THE UNIVERSITY OF EDINBURGH

IMPROVING INSTRUCTION IN SOCIAL WORK: TWO EVALUATION PARADIGMS

A THESIS SUBMITTED TO THE FACULTY OF SOCIAL SCIENCES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR
THE DEGREE OF Ph.D.

DEPARTMENT OF SOCIAL ADMINISTRATION

EDINBURGH

FALL, 1978

DECLARATION

I Kenneth H. Gordon do declare that the thesis entitled, "Improving Instruction in Social Work: Two Evaluation Paradigms" was composed entirely by myself and is my own work.

Dated June 30, 1978

UNIVERSITY OF EDINBURGH

FACULTY OF SOCIAL SCIENCES

The undersigned certify that they have read, and recommend to the Faculty of Social Sciences, for acceptance, a thesis entitled, "Improving Instruction in Social Work: Two Evaluation Paradigms, submitted by Kenneth H. Gordon in partial fulfilment of the requirements for the degree of Ph.D.

. . . . . .
Supervisor

. . . . . .

. . . . . .

Date . . . . . .

UNIVERSITY OF EDINBURGH

ACCESS TO A THESIS and PUBLICATION OF ABSTRACT

[The following Declaration is designed to protect the Candidate while expediting the availability of his work to other scholars. Candidates should note that the thesis may be consulted by registered users in Edinburgh University Library or in the libraries of other educational establishments but may not be issued for loan outside any library to individuals except academic staff of the University of Edinburgh. All applicants for consultation, loan or photocopy of the thesis have to sign an appropriate declaration.]

Full Names of Candidate (surname first) GORDON, KENNETH HUNTLY

Candidate for degree of Ph.D. Faculty SOCIAL SCIENCES

Title of Thesis IMPROVING INSTRUCTION IN SOCIAL WORK: TWO EVALUATION PARADIGMS

1. I hereby authorise the University of Edinburgh to publish the abstract of this thesis, and to authorise others to do so, for scholarly purposes and with proper acknowledgement of authorship.

Please DELETE two paragraphs of 2–4 below, leave one of them undeleted, and sign the form. Save in exceptional circumstances, PARAGRAPH 2 is recommended.

2. I hereby give my permission to Edinburgh University Library to photocopy this thesis in whole or in part for use by individuals, on condition that each person for whom a copy is made signs the following declaration (which will be preserved in Edinburgh University Library and available there for inspection by me);

   'I undertake fully to observe the author's copyright in this thesis, not to publish the whole or any part of it without the author's written permission, and not to allow any other person to use the copy made for me'.

OR

3. ~~I wish all applicants for photocopies of the whole or any part of this thesis to be referred to~~ me individually; but I recognise that if Edinburgh University Library ~~cannot~~ by reasonable enquiry ~~discover my~~ address the University Librarian ~~will decide~~, after consulting my Head of Department, whether or not ~~to make~~ a photocopy on obtaining from the applicant a declaration as in 2 above. Please delete the following sentence if desired: The bibliography/ ~~list of works consulted may be photocopied providing the applicant completes a declaration~~ as in 2 above.

OR/

ABSTRACT OF THESIS

Name of Candidate GORDON, Kenneth Huntly

Address 11239 - 75th Avenue, Edmonton, Alberta, Canada

Degree Ph.D. *Date* June 30, 1978

Title of Thesis IMPROVING INSTRUCTION IN SOCIAL WORK: TWO EVALUATION PARADIGMS

This dissertation examines and contrasts two dissimilar investigative paradigms which may be applied to the systematic evaluation of instructional sequences for social workers. The primary orientation is toward evaluation aimed at generating improvements in instruction. Systematic evaluation is seen to be a necessary process in rational decision-making. Such investigations are always guided by a research paradigm of some sort. The choice of paradigms, conscious or not, largely determines the sorts of information produced. An overview of the evaluation of instructional programs includes: definitional elements, distinctions, and issues; purposes and foci associated with such activity; information sources and media of data generation; and aspects of instrument and information quality. Review of the relevant literature discovers a dearth of published evaluation studies in social work education. It further indicates that one orientation to evaluation research has dominated the scene.

The two paradigms under study are described and contrasted in theoretical terms. Critical comments and arguments between paradigmic proponents are documented. The "classico-experimental" orientation is seen to stress assessments of effectiveness; achieved through operational definition of instructional objectives and application of experimental or quasi-experimental procedures. Typically, the resulting information is quantified. Emphasis is on measurement and prediction. In contrast, the "socio-anthropological" approach is perceived to be more concerned with

Use this side only

ABSTRACT OF THESIS

Name of Candidate GORDON, Kenneth Huntly

Address 11239 - 75th Avenue, Edmonton, Alberta, Canada

Degree Ph.D. *Date* June 30, 1978

Title of Thesis IMPROVING INSTRUCTION IN SOCIAL WORK: TWO EVALUATION PARADIGMS

exploration, description, and interpretation of the instructional sequence as a whole. Quantification is not necessarily prized and a wide variety of data gathering media are used to progressively focus on significant phenomena which emerge during the course of the study. Emphasis is on tailoring investigative activity to information requirements and characteristics of the milieu examined.

These paradigms are applied in evaluating two very similar short courses for social workers, and reports are presented. Attending to experience in those studies, the application and the paradigms are further contrasted. Observations on differential reliability, validity, generalizability, and utility are noted. A cursory exploration of student rating scales, as evaluation devices, is also included. Tentative conclusions with regard to paradigmic strengths, weaknesses, and appropriate applications are drawn. Suggested practice principles are itemized and the feasibility of paradigmic synergy is examined. It is concluded that, while a true amalgamation is not feasible, comprehensive evaluation studies should draw upon the strengths of both paradigms. Some implications for instructional planning and a new approach to curriculum revision in social work education are explored. A number of major requirements for further research are outlined. Though selective, the Bibliography documents numerous sources.

PREFACE

I applied to study at Edinburgh University with the rather vague intention of examining methods by which at least certain forms of social work education might be evaluated. This preoccupation had arisen from feelings that, as a training officer in a large government social welfare agency and as an instructor on a community college course for social service aides, I was really unable to put my finger on many specific changes which my ministrations had induced in the students and trainees who suffered under my tutorship. Some of my colleagues had expressed similar concerns. Society, through one agency or another, had delegated to us responsibility for the "education", "instruction", or "training" of individuals who were to be employed in the provision of certain social services. We "instructors", "teachers", or "trainers" had been "instructed", "taught", or "trained" to take on professional roles in social work, psychology, psychiatry, sociology, or social administration, and those experiences--combined with our practice experience--led us to believe that we were at least partially prepared to assume responsibility for teaching others. In my own case, I duly created course curricula and "taught" as best I could, while at the same time trying to learn all that I could about teaching as an activity in itself. It was not until after I had been immersed in this pursuit for some time that I began to wonder just how much of what I "taught" was being learned by my students, whether

what I was teaching was relevant to the future needs of my students, and if there might be some more systematic means of improving my efforts at instruction.

I had read a number of books and articles which dealt with instructional methods and curriculum development, and had attended workshops, seminars and the like which were designed to teach one how to teach. Both of these sources of information had brought me to the conclusion that questions about improving my instruction might be answerable if I could find reliable, systematic methods for evaluating instructional activities.

Once accepted at Edinburgh University, and after an initial period of more closely defining my concerns and negotiating with my study supervisors, I formulated an intention to study "methods for the evaluation of in-service training programs for social workers." This seemed to be an apt topic, as the bulk of my teaching experience had been in providing "in-service" and "continuing" training to social work personnel. Furthermore, the area of study had to be limited in some definite ways and the short-term nature of in-service training not only articulated with my interests but made exploratory projects practicable. I launched myself into locating literature on the evaluation of educational programs and sought out individuals and groups who might inform me further, as well as agencies that might offer me possibilities for the study of in-service training programs. The reading led me into the deep and murky waters of educational assessment, curriculum

evaluation, and instructional technology. The contact with social work instructors and trainers led me to conclude that very few in-service training programs would be available for study. Local Authority employees in Scotland were preoccupied with an overall reorganization and their in-service training programs had been, for the most part, temporarily suspended. It became apparent that if I wished to find instructional sequences to study I would have to adopt a somewhat different focus. The university courses, particularly those at Edinburgh University, appeared to offer possibilities but they also presented practical drawbacks which forced me to rule them out. Generally, university courses were too long for my purposes. I had limited time and resources at my disposal, and would not have been able to try out more than one evaluation method unless much shorter courses became available. I knew that evaluation studies tended to require much more time than that spent simply in attending the instructional program. Furthermore, I was still primarily interested in evaluating short courses of one form or another.

By this time I had learned that there was a third sphere within which social workers in Britain were given instruction. Forms of education and training were provided through the programs operated by three organizations: The Central Council for Education and Training in Social Work (CCETSW), The National Institute for Social Work (NISW), and The London Boroughs Training Committee (LBTC). Each of these organizations contributed to the continuing education

of Britain's professional social workers. The CCETSW fulfilled quite a number of functions, only one of which was to meet ongoing instructional needs of practicing social work personnel. Furthermore, it had recently opened a new office in Edinburgh in which staff were attempting to develop a number of short courses for social workers in Scotland, Northern Ireland, and the North of England. The NISW was based in London and was also involved in fulfilling a number of diverse functions. One of its primary activities was the development and presentation of short courses in social work. The LBTC had been formed to meet some of the training needs of social workers employed by a number of the London Boroughs' social work departments. Its major function, then, was the operation of training courses for social workers, and its course program appeared to be a promising source of material for my studies.

My contacts with representatives of these organizations indicated that each was willing to provide me with some assistance. All had shown initial willingness in allowing me access to their courses. My focus, therefore, changed slightly, from "in-service training courses" to "short courses of continuing education." However, I was not yet ready to propose a full-blown study. My reading had led me to believe that there was one major approach to the evaluation of instructional programs--the approach which I shall characterize later as the "classico-experimental" paradigm. In consultation with my study supervisors, I decided that

the best way to learn about the strengths and weaknesses of the paradigm was to try out a simple model of it in evaluating a course. Consequently, I searched the course programs of the three previously mentioned agencies and decided that a short course offered by the CCETSW, entitled "Caring for the Elderly", best met my needs. It appeared to be a straightforward course designed to impart to social workers various concepts and attitudes which would supposedly assist them in meeting the needs of elderly clients. My major reasons for choosing it were three-fold: (1) my reading had led me to the opinion that evaluating courses which were primarily aimed at the development of cognitive capabilities in students was relatively less difficult, and this course was obviously very heavily oriented to knowledge-giving; (2) the CCETSW and Age Concern Scotland who were sponsoring the course appeared to be very willing to support and assist my evaluation; and (3) the course was short.

I attempted to keep up with my reading in evaluation methodology while at the same time developing plans and instruments which could be used to evaluate "Caring for the Elderly." This was a very instructive process in that I experienced first-hand some of the difficulties about which I had only read. Eventually, I was as prepared as circumstances would permit. I had created a comprehensive knowledge test and an attitude scale which were directly related to the subject matter of the course, and had prepared a student information sheet and various student

evaluation rating scales for administration. It had been impossible to find a control group against which to compare performance of the students, but in every other way I had tried to adhere to a quasi-experimental approach. I administered tests at appropriate times, attended the course, collected, collated and analysed my data, and eventually produced an evaluation report which was well received by the CCETSW and Age Concern Scotland. The report represented a serious attempt to systematically evaluate a social work course, and it appeared to be of some practical use to the organizing agencies. If nothing more, it made some useful recommendations for the operation of future courses and promoted a closer dialogue between Age Concern Scotland and the CCETSW. Furthermore, I had found that involvement in every aspect of the project had given me a first-hand look at the multitude of problems and issues which arise in evaluating a program of instruction in social work. Unfortunately, it seemed that there were so many problems and issues--many of which will be discussed later in this dissertation--that in my eyes the study had been an utter failure. It had not begun to answer many of my questions about the instructional program and the validity and reliability of much of my "evidence" was highly questionable. My analysis suggested that limitations on the quality of my data were partially due to my own inexperience and extreme lack of resources, but it also suggested that the whole quasi-experimental approach might be of questionable utility.

I had, by this time, read enough to realize that there were not only a large number of approaches to evaluation which fit relatively comfortably under the experimental orientation, but also that there was at least one other major research paradigm which might guide evaluative efforts. This second orientation had been characterized as the "socio-anthropological" paradigm and was obviously a major departure from the traditional plan of attack in evaluating instructional programs. I came to the conclusion that this approach to evaluation should also be given a trial on an instructional sequence for social workers. I later learned that it was currently being applied in another study concerned with elements of social work education at Edinburgh University.

It seemed to me at that time, and still does, that one should examine his tools before attempting to use them. One tool may be of particular use in a certain job, while another is not. The second, however, may be much more useful for another sort of job. I viewed the two evaluative paradigms as tools which might be used to do the job of evaluating programs in social work instruction, and I felt that those paradigms had to be closely assessed before one could make decisions about applying them. In effect, I decided to begin a comparative analysis of the two research paradigms as they applied to evaluating short courses in continuing education for social workers. In order to do so, I set out to find a situation within which both might be tried out under very similar conditions. I hoped that

trials under similar conditions might at least point out more clearly the paradigms' differences and similarities and, perhaps, might generate some information on their respective utilities. Obviously, applying them in similar circumstances was likely to give me some information on the utility of each within a particular sort of situation. It would not, however, provide me with a great deal of evidence about how each might be used in very different situations. That sort of information could only be gathered from their application in a wide variety of different circumstances. I had very few of the resources necessary to undertaking such a broad project. I resolved, therefore, to limit my studies to one rather circumscribed set of circumstances which appeared to be fairly representative of short courses for social workers.

I set out three major conditions or characteristics which would have to be present in the milieu I visualized for the study: (1) subject matter of the course had to be highly germane to social work education in general; (2) the instructional programs had to be short enough and temporally far enough removed from each other to allow for preparation before and analysis after each application; and (3) teaching methods had to be relatively representative of those used by many social work instructors. The first and third characteristics were needed to allow for at least a degree of generalizability of my findings; the second, for practical reasons related to my own resource limitations and because of my interest in short courses of instruction.

Only one locus of study which displayed these three characteristics presented itself. The NISW was planning to present a short course entitled "A Unitary Approach to Social Work Practice" on two occasions--once in January 1976, and again in May 1976. These proposed programs appeared to meet my criteria for acceptance. Furthermore, discussion with the course organizers had opened the door to evaluation of both programs. I resolved to proceed with the study.

Throughout the project NISW staff remained entirely supportive of my efforts. Since, in my pilot study, I had already used the "classico-experimental" paradigm, and because of certain characteristics of the two paradigms, I decided to apply the "socio-anthropological" model to the January program and the "classico-experimental" model to the May run. The details of these two applications are included in later sections of this dissertation, along with the evaluation reports and discussions of issues which arose at the time.

The final phase of my investigation had necessarily to be some form of contrasting between the two paradigms as they were applied in these specific instances, and as they might be applied in other situations within the sphere of social work education. I have attempted not only to examine some of the issues, problems, strengths, and weaknesses associated with using both paradigms, but also have tried to develop a few practice principles which may help to guide and instruct others who wish, through evaluation, to improve

social work instruction. Some of those guidelines and principles arise directly from the comparative study of applications; others flow more from the relatively extensive reading which has been necessary for me to understand and use the paradigms. Finally, I have ended the dissertation by suggesting areas requiring further study and by proposing a broad curriculum planning orientation related to the investigative paradigms.

In effect, then, I have made a fairly extensive, and hopefully not overly-destructive, foray into another domain--that of educational evaluation--in order to bring back information which could lead to the eventual improvement of instructional programs in social work and raise the level of discussion on evaluative issues in social work education from its present abysm. In so doing, I have attempted--in a limited way--to test out two major conceptualizations of evaluation methodology in a social work context.

May 1978 Kenneth H. Gordon

ACKNOWLEDGEMENTS

I shall never be able to repay my indebtedness to those who have supported, challenged, and stimulated me in the preparation of this dissertation. One does not undergo such a period of study thousands of miles from home without many doubts and difficulties. In that regard, I should like to thank the faculty of the Department of Social Administration, particularly Dr. John Triseliotis and Mr. Neil Frazer. Without the benefit of their expertise and emotional support I should never have learned so much in such a short time. Without their critical advice, even though sometimes unheeded, maintaining an overall sense of balance and perspective would have been very difficult. In that same regard, I should like to express deepest gratitude to my parents. Without their unconditional support, I would never have finished.

I should also like to acknowledge the assistance of Miss Anne Vickery, the students and staff of the National Institute for Social Work, the Central Council for Education and Training in Social Work, and the London Boroughs Training Committee. Their generousity, acceptance, and willingness to undergo close scrutiny made the entire study possible. The financial support of National Health and Welfare Canada and the typing skills of Mrs. Marilyn Wahl were also indispensable.

Finally, to Kay, to Merle and Tom, and to all the lovely people of Scotland who made my sojourn in their

beautiful land an exciting and rewarding phase of my life--thank you.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

THE EVALUATION OF INSTRUCTION: AN OVERVIEW

The Purpose of Evaluation

When one pauses to consider, there are obviously many reasons for wishing to undertake an evaluative investigation of a program of instruction. We may be interested, for example, in deciding: whether to continue or discontinue the program, whether to add or drop specific program strategies and techniques, whether to institute similar programs elsewhere, how to allocate limited resources amongst competing programs or program elements, whether to accept or reject a program approach or theory, upon means of improving program practices or procedures. The primary emphasis in this dissertation will be upon the evaluation of instructional programs in social work which leads to decision-making about program improvement. At a higher level, the aim is toward seeking improvement in student learning. However, although this more "ultimate" goal may at first appear to be simple enough, it must be recognized from the outset that there is precious little empirical evidence to support the assumption that making apparently desirable changes in instructional programs will necessarily lead to improvements in student learning, or that improvements in student learning will eventually lead to improved services for social work clients. In their everyday affairs, most social work instructors, educators, trainers, and

teachers appear to take such assumptions for granted. That they should not be taken for granted may become more clear as this dissertation unfolds. For now, they should serve as a warning that we are entering a very complex area of enquiry which is, of necessity, rife with conflicting opinions, controversial issues, and extremely difficult methodological problems.

We might begin by taking note of the fact that evaluation studies can be, and relatively often are, undertaken for reasons which have very little to do with our proposed goal of assisting student learning. It may transpire, for instance, that a university faculty member will decide to undertake the "evaluation" of a sequence of instruction simply because he or she feels a need to produce a written article. Exactly what he produces may not seem to matter much, so long as it appears to be respectable and relevant to his "locus of expertise". Particularly in professional schools where a premium is placed on "applied research", and especially in schools and departments where there is intense pressure on faculty to produce published material, we can expect to find at least a few "evaluation studies" which are primarily aimed at meeting this expectation. Similar activity and motivation can be found, perhaps with even greater frequency, in governmental and other organizations because, for one reason or another, "the boss" wants a paper or a study. The resulting product can be beautifully illustrated with statistical tables, graphs, and flow charts of exceeding complexity. With the

help of cooperative students, a little time to spare, and a modicum of writing skill, the expectation to produce an article can be met and the producer may have taken one more small step up the academic or organizational ladder. It is also entirely possible that the article or report will be of absolutely no use to anyone but its producer. One must not, however, automatically assume that this will be the case. Such studies may, almost inadvertently, provide information and pose questions which prove to be very useful in another context. The point here is that such purposes--and the products to match--do exist, and that they represent only one example of the wide range of reasons for which people undertake evaluations of instructional programs. As Weiss notes, evaluations are ideally undertaken because answers are being sought to pressing questions about program future, but evaluations are also undertaken for other reasons; e.g., to delay a decision, to cover a pre-ordained decision in the mantle of evaluative research, to generate support, to fulfill grant requirements.[1]

If we concentrate more closely on studies <u>designed to evaluate</u> instructional programs we may see that they are aimed at the development of knowledge: knowledge which, it is hoped, will be of special utility to people who must make some sort of decision(s) about the instructional program in question. It is generally assumed, and relatively well substantiated, that when people must make a decision about some aspect of human activity, they can usefully equip themselves for the task by gaining relevant knowledge

and coming to an understanding of the activity. In order to do that, information must be available. The ascertainment of certain pieces of information and the discovery of relationships between those pieces of information, in the service of human understanding and decision-making, comprises a major mission for evaluation studies.

There are, in fact, two general decision-making functions which evaluation can help to fulfill: decision-making to effect improvement, and decision-making about accountability. Some writers refer to the first area as "decision-making" and the second as "accountability."[2] The process of assessing accountability, however, involves decision-making just as much as does the process of effecting improvements. In both, decision alternatives must be judged, and for both, information must be produced. These functions, then, are not as clearly distinguishable as they might initially appear to be. Michael Scriven has noted that the roles of evaluation are enormously varied but that, when analysed, they divide into two classes: formative and summative.[3] Formative evaluation is an activity which assists in developing curricula. It is part of the curriculum development process in that it provides feedback to assist in the development of an instructional product. As such, it must address questions about content validity, vocabulary level, utility, appropriateness of media, efficiency, and so on. On the whole, it may be viewed as evaluation internal to the program and aimed at improvement of the program. Summative evaluation addresses questions about

the effectiveness of the program once it is producing a product. Further, it often aims to determine whether the program is more cost-effective than its competitors. It is very often external to the instructional program and provides consumers, or potential consumers, of the program's products with assessments of the merit of those products. The concept of formative evaluation correlates very closely with the function which we have identified as "decision-making to effect improvement"; the concept of summative evaluation is more often associated with the function of "decision-making about accountability". Although there are slight conceptual differences between the members of these sets, generally, we can use one to inform, embellish, and approximate the other. Evaluation may be viewed as "formative" if it is conducted proactively to serve decision-making about program improvement, and "summative" if it is used retroactively to serve accountability considerations.[4] Formative evaluation, however, may serve to inform the summative evaluation process and summative evaluation may, at times, be viewed as part of a larger formative evaluation process. For instance, a formative evaluation of instructional methods used in a course may provide valuable information to the summative evaluation of that course. The summative evaluation of that course may form part of a formative evaluation of the entire departmental program within which that particular course operated. So, although it provides a useful conceptual tool, the formative-summative distinction (as well as the improvement-accountability

distinction) is not entirely clear-cut and might be more usefully viewed as a continuum, rather than a dichotomy. Formative can feed summative; "improvement data" can feed "accountability data". Summative can inform formative; "accountability data" can feed "improvement data."

Thus far we have been content with a conception of evaluation which includes only information production. A more complete definition, however, stresses that it is an activity of appraisal; of finding the value of something. Dictionary definitions, for example, equate evalution with appraisal and assignation of value. Value judgements are integral to evaluation, not only with regard to the subject of an evaluation but with regard to evalutive means as well. A more complete definition, then, stresses that evaluation is not simply an activity aimed at producing information, it is also an activity which is necessarily bound up with the making of value judgements. In their everyday lives, people constantly make informal evaluations of other people, of the behavior of other people, of concepts, and of "things". For instance, the prospective male client with the opportunity to choose a social worker may note that one worker is female, young, and physically attractive to him, while another is male, older and physically unattractive to him; for whatever reasons, he decides that he would prefer to see the former. In so doing he has evaluated, albeit in a rather limited and informal manner. He has used information and made value judgements to arrive at a decision. In similar ways, he may arrive at decisions about how he will deport himself with the

social worker, how he will present his problem, how he will leave the office, and how he will spend the money which she has provided. If he was an extremely fastidious man, he might seek much further for information, judge it differently, and adopt different alternatives at any or all of these decision points. He might, in fact, formalize the evaluation process in an attempt to more systematically arrive at the best decision. The more important a decision is, the more important that the evaluation method become systematically formalized.

With regard to social work instruction, evaluations will be carried out--whether with a high degree of formality, or none at all. If, however it is hoped that the resulting decisions will be good ones--based on reliable, valid, and useful information--then attempts to formalize or systematize such evaluations must be made. Furthermore, just as it is to be hoped that those having to make decisions about instructional programs will use reliable and valid information in their deliberations, it is also to be hoped that those who produce the information will do their utmost to ensure that it is valid, reliable, and useful to the decision-maker. This is not intended to suggest that the information-gathering and the decision-making agency might not reside within the same organization, collective, or person. Circumstances abound in which both segments of the evaluative function reside within one person or group; but, at the same time, there are many circumstances where decision-makers delegate or commission other persons or agencies to provide them with

relevant information. The basic evaluation process is theoretically the same in either case.

We have concluded that evaluation may be viewed as an attempt to systematically develop information for judging decision alternatives. A distinction must now be made between "validation"--attempting to ascertain whether an instructional program has achieved its aims--and "evaluation"--assessment of the value of the instructional program. This is not the same distinction as that between summative and formative evaluation. Validation seeks to assess the extent to which an instructional program has been successful in teaching what it set out to teach (internal validation) and whether the aim of the program itself was realistically based on needs external to the program, such as the needs of students (external validation). Evaluation, on the other hand, attempts to assess the overall value of the program, not just the achievement of its aims. Obviously, if our approach to evaluation is to be systematic, validation is an important component; but it should also be recognized that validation does not comprise the whole of evaluation. Validation is concerned with assessing effectiveness (Has the program reached its stated objectives?); evaluation is interested in all effects, and in the relationship of effects to costs (Was the program worthwhile?). Although this distinction may appear to be somewhat trivial at the moment, it assumes much greater importance as we examine evaluation theory and methodology in greater depth.

Another distinction which should be noted at this time is between evaluation activities which seek to appraise

the qualities of elements within an instructional program (intrinsic evaluation) and activities which are concerned with the program's effects on students (payoff evaluation).[5] "Intrinsic" evaluation activities seek to establish the nature of the program (What is it?), while "payoff" evaluation activities seek to establish the program's effects (What does it do?). Although intrinsic evaluation is important (we need to know <u>what</u> we are evaluating), the evaluation is not complete unless we determine and judge the effects of the program. Again, this is not the formative-summative distinction. Intrinsic evaluation activities can serve both formative and summative evaluation functions; so can payoff evaluations. Again, the distinction will become more important as this dissertation proceeds.

In order to provide decision-makers with useful information for judging decision alternatives, it is, of course, necessary to obtain that information. As already suggested, much evaluation activity is aimed at doing just that. But in order to do that well, we must first delineate the sorts of information which we most wish to obtain and present. Information is not likely to be either useful or available if we do not first delineate the sort of information required. Delineating, obtaining, and providing information are the three steps. Delineating questions to be answered and providing information are activities which require the presence of, and interface between, data producer and decision-maker; obtaining information is a technical activity involving observation, measurement, and data

processing which is executed mainly by the information producer. By incorporating these steps into a definition of evaluation, we arrive at the following: "Evaluation is the process of delineating, obtaining, and providing information for judging decision alternatives."[6] If this is broadened to include judging activity within the evaluation framework, which, as we noted earlier fits with our more complete definition, we arrive at something like: "Evaluation is the process of delineating, obtaining, and providing information and making judgements useful for selecting decision alternatives." Although there will be objections to any definition of evaluation, this one appears to: (1) be relatively comprehensive in theoretical terms; (2) allows the argument that evaluators must judge instruments, weightings, and the selection of criteria; and (3) encompasses a very large majority of the evaluation work actually being carried out in the field.

In order to gain a more explicit conceptualization of the various purposes which guide evaluations of instructional programs, it may be worthwhile to look at a few of the people involved with the instructional system and examine one or two of the sorts of decisions which they must make. Although a complete and detailed list of participants in, and associates of, any particular instructional program in social work might be very long, for our purposes they may be categorized as: "instructor", "student", "administrator", and "outside agencies".

An important stimulus to my own interest in evaluation arose from feelings that, as an instructor, I did not have much information about the effects my instructional efforts were having on students. Since it was my desire to become a more effective teacher, I needed a means of acquiring more and better information. Other teachers, instructors, trainers, and educators have undoubtedly had similar desires and needs. They must make a variety of decisions about the programs in which they are involved. They will require internal validation information to assess the extent to which they are effectively achieving their objectives, and information on how those objectives might be more effectively achieved. They will require external validation information to assess the extent to which those objectives meet external needs, and information on how those objectives might be profitably altered. They will require information on the overall impact of their instruction, both for "improvement" and "accountability" purposes.

In order to meet the formative function--to improve instruction--instructors will have to make decisions about subject matter, instructional materials, teaching methods, physical circumstances, social environments, and so on. An instructor will have to decide whether to continue on his present tack, terminate, or alter it. He may need to make decisions about students; which to accept or reject, which are most likely to benefit in circumstances which are largely not amenable to change, whether they are ready to learn what he intends to teach, or whether what he intends to teach is appropriate to their needs and desires. The

list could go on indefinitely and would vary as the teacher's purposes and circumstances varied. In this paper we are making the basic assumption that one of the instructor's major purposes is to promote learning in students. That would appear to be the primary function assigned to him by society and which he accepted when he took on the mantle of teacher. To the extent that he wishes to maximize student learning, the instructor will attempt to influence into existence an optimal set of circumstances for that learning. In order to do that, he will wish to know about the effects and effectiveness of his instructional procedures. materials, theoretical formulations, and so on. In short, he will wish to evaluate.

The student too must make decisions which will vary with his purposes. In some cases he might be concerned about the effectiveness of instruction; either of programs in which he is participating or programs in which he may participate at some future date. It is likely that he will be even more interested in assessing some of the other effects, such as vocational and social outcomes, which appear to flow from program participation. If his primary concern is with receiving some sort of credit, grade, certificate, or degree, he may wish to choose a course or set of courses which will provide the opportunity or increase the probability of his reaching that goal. If he is more concerned with having an experience which holds his interest, he may choose courses which appear likely to provide opportunities that interest him. Alternatively, if

he is concerned with preparing himself for certain future activities, such as social work practice, he will wish to choose courses which will assist him in that preparation. Of course, there are situations in which he would seem to have little choice in this latter regard. If he commits himself to the process of professional education in social work, for example, he will be required to undertake certain courses of studies. However, he will still have decisions to make: how best to allocate his time and effort, how best to prepare for a particular role which he envisons for himself, when it is best to accede to the requirements placed on him and when it is better to attempt alterations. Again, we make the assumption that the social work student's major purpose for involving himself in an instructional system is to learn. Again, we note that this is not likely to be his only purpose, but assume that it is ranked high in his personal hierarchy of values. To the extent that the student wishes to maximize his learning, he will attempt to influence into existence an optimal set of circumstances for that learning. He too will wish to evaluate in order to arrive at the best possible decision alternatives.

Those who administer to the operations of schools, universities, training authorities, or parts thereof, will also have to make decisions, the forms of which will vary with their purposes. Here the need for summative evaluation may become more apparent. In addition to requiring information to improve system effectiveness, the administrator must deal with accountability demands which arise from

outside of the instructional program. The university administrator, for example, may have to make economic decisions about the allocation of limited departmental funds. He may also have to make personnel decisions; about the selection, deployment, tenure, pay, or rank of faculty; or decisions about the selection and retention of students. Whether he be a departmental administrator, a university vice-chancellor, a head-teacher, or the director of a training authority, the administrator will have to make decisions about instructional programs, some of which will be concerned primarily with program improvement and some concerned with meeting accountability demands. Clearly, he too will need to evaluate.

The term "outside agencies" is used here to represent a broad range of organizations, individuals, and collectives that have a stake in our hypothetical instructional program. It could include such diverse groups as: agencies that employ, or may wish to employ, students in the program; professional bodies; other training or educational organizations; other instructors; suppliers of curriculum materials; agencies providing practice opportunities for students; and even prospective consumers of program products, like social work clients. Clearly, this is an extremely broad category within which values and purposes will vary tremendously. An employing agency might wish students graduating from the instructional program to possess certain skills which are consonant with the agency's functions and needs. Another educational or training organization might

wish the instructional program to articulate very closely with its own--or not at all. A particular collective of prospective clients might wish the graduates of the program to possess knowledge which is relevant to their own special needs and desires. Obviously, it is from within this very broad range of interests that demands for accountability will most often originate. Accountability demands and conflicting interests will occur at every decision-making level, but it is in this wider environment--where so many groups with conflicting ideals, desires, needs, and interests exist--that the demands and value conflicts will be most manifest. To the extent that any of these individuals, organizations, or collectives wish to maximize their own benefits from the instructional program, they will wish to evaluate it.

We may draw the conclusion that instructional programs are attended by a variety of individuals and groups who must make decisions about them. They must make decisions about the effects or effectiveness of the program as a whole, or about components of the program. They must make decisions about ways in which it should be altered; they must make comparisons between programs or components of programs, and they must make assessments of accountability. In addition, each decision-maker assesses alternatives with regard to a particular value orientation, and these value orientations tend to vary considerably. Evaluation activities designed to meet the information requirements of any of these "audiences" are not likely to meet the requirements of many

others. No evaluation, not even the very best, will provide information on all questions and concerns that people will think of. In fact, the purposes of one interested group may be incompatible with the purposes of others. For this reason, a complete model of evaluation would provide guidelines for the identification of "audiences", as well as assessment of their unique and common information needs.

Focus of Evaluation

In order that we may more closely examine some of the methods which might be used in evaluating instructional programs, we must concentrate more specifically on the objects of those evaluations. What, exactly is it that people have to make those decisions about? In other words, what is the focus of this evaluation activity? Thus far I have used the terms "instruction" or "teaching" almost exclusively and have intentionally limited use of the terms "education" and "training". Those latter terms are generally defined somewhat differently. Whereas the word "training" is commonly used to denote an activity aimed at the development of a skill or set of skills, "education" denotes an activity aimed more broadly at the development of intellectual and ethical faculties. More specifically, a person may be trained to do something, while he may be educated to learn to do something. There are many distinctions drawn between education and training and, although further examination of them might prove interesting and instructive, our focus upon instruction in social work is intended to

encompass both. The enterprise which we commonly refer to as "social work education", "social work training", or "social work instruction" clearly has both an educative and a training function. It aims at the development of some constellation of knowledge, values, and skills in its students. For that reason we will continue to use the broader terms "instruction" or "teaching" to denote all such efforts to develop knowledge, values, and/or skills in learners.

The term "program" has been used to indicate an agenda of activities which are or will be carried out. Perhaps a better term is "process", indicating a whole series of continuous actions by which something is accomplished. We might also consider using the word "system", which denotes a methodically arranged set of ideas, principles, methods, and procedures. The terms instructional or teaching "program", "process", or "system" will be used interchangeably to indicate a whole systematically arranged set of ideas, principles, methods, and procedures, by the use of which instruction or teaching is carried out. In so doing, we impose no limitations on the size or complexity of the program. Programs may vary on a number of dimensions (e.g. scope, size, duration, clarity and specificity of input, complexity and specificity of goals, innovativeness) but each constellation of activity may be regarded as a "program" in its own right.[7] The terms, then, may be used in referring to the very large and complex network of classes, seminars, tutorials, demonstrations, and practice opportunities

which make up the "instructional program" of an entire graduate school of social work. At the other extreme, they may be used in referring to one discrete sequence of activity in one particular classroom over a very short period of time. Although the procedures used to evaluate these two programs may vary considerably, the basic theoretical conceptualization guiding those procedures does not. Evaluation techniques must be developed with reference to the particular program(s) to be evaluated; basic theoretical conceptualizations of evaluation are as valid for one as another.

We have noted that instructional programs may be seen as "systems". That viewpoint may be further used to assist us in organizing our enquiries into the nature of evaluation.[8] For example, a "systems" orientation suggests that we might view instructional activities in terms of the inputs to the program, the processes which occur within the program, the outputs of the program, and the environment or context within which the program occurs. Since we are primarily interested in the nature of the instructional system itself, we will concentrate on inputs, processes, and outputs; but it would be virtually impossible--and naive in the extreme--to discuss any system without occasionally paying attention to the environment within which it operates.

Inputs to an instructional program include personal characteristics of people, ideas, principles, intentions, beliefs, proposed methods and procedures, materials, information; whatever the participants or environment contribute to it. As such, the concept appears vague and rather useless

but, perhaps, it can be at least partially clarified by examining a few examples of systemic inputs which originate from diverse sources.

The instructor or teacher involved in planning and operating a program obviously provides some of these inputs. His relevant knowledge, choice of goals, personality characteristics, communicatory abilities, relevant beliefs and values, ability to motivate and stimulate, and so forth, will all--in one way or another--provide material out of which the instructional program will be structured. In fact, he is likely to be a highly significant source of input and many studies of evaluation methodology, as well as many evaluation studies themselves, place a heavy emphasis on instructor-related inputs.[9]

It is just as obvious that students, individually and collectively, contribute a great deal to determining the structure of the instructional process. Student-related variables such as personality characteristics, group characteristics, "readiness" for learning, educational and employment experience, relevant knowledge, values and beliefs, and social characteristics, all represent resources used in the instructional system. However, though they often play a major part in structuring the system, they are often disregarded or treated as relevant only in terms of student "assessment", rather than student "contribution". Some instructors, and a few very naive evaluators, behave as if the student enters the instructional system tabla rasa. Later we will see how disregard for student-based characteristics can invalidate an instructional program and/or any evaluation of it.

A third major source of input to the instructional system is the teaching plan and instructional materials brought to it--usually by the instructor. Any determination of a plan or set of materials to be used in the program will, of course, assist in structuring it. A number of studies have concentrated on evaluating and investigating text books, programmed learning materials, expository methods, and teaching aids.[10] Input variables such as these might be evaluated in terms of their relevance, palatability, ability to motivate and stimulate, communication characteristics, and so on.

Clearly, the environment within which the instruction takes place may provide very important inputs to the program. One might consider a multitude of environmental variables (social, physical, political, philosophical, organizational) which may have both overt and covert effects on the instructional program.[11] In particular, an evaluator might wish to pay attention to inputs arising from the social and physical environments.[12]

The processes occurring within our instructional system may be viewed as existing on two dimensions: intrapersonal processes and interpersonal processes. The intrapersonal or "internal" processes are activities imputed to the student and are characterized by concepts such as knowing, comprehending, applying, analyzing, synthesizing, evaluating, receiving, valuing, organizing, and characterizing.[13] Various conceptualizations of intellectual structure view these processes in different ways.[14] Two major taxonomies

of educational objectives have used them as organizing concepts.[15] The field of educational psychology is replete with discussions of such concepts and some of them have become almost universally recognized. There is, however, a basic difficulty in studying these processes: we are able to directly perceive what people do but still have rather limited means for infering intrapsychic states and processes. Interpersonal processes, however, are more clearly open to observation. Those that would appear to concern us most in the present context are those occurring between instructor and student, instructor and instructor, and student and student; but there will undoubtedly be situations in which administrator involvement is of at least equal importance. Further, we will observe numerous interactions occurring between various elements in the instructional system and elements in its environment which are important and, in some circumstances, may become crucial process variables requiring thorough examination in the course of an evaluation. It is, however, in the interactions between instructors and students, and students and students that we will be able to most clearly view many of the most crucial interpersonal processes. We may observe an instructor lecturing to an amphitheatre full of students, we may hear students discussing relevant issues over a cup of coffee, we may watch a group of students working on a prescribed learning task, we may see an instructor demonstrating the correct use of a particular tool or technique. In each of these activities we will be witnessing an

interactional process which might be expected to bring about learning in students, and might be characterized as "operationalizations" or "implementations" in instructional programs.

The outputs of an instructional program might also be called "outcomes", "effects", or "consequences" and are, as was intimated earlier, tremendously varied. The related term "impact" is more commonly used to denote further consequences of "outcomes"; often larger or more long-range effects. Our primary locus of interest in this dissertation is in student learning outputs, but we have also recognized that an instructional system might be seen to produce many effects which are not directly connected with student learning. A school, for example, which operates what is generally perceived to be a "successful" educational program might deepen or broaden its ties with local graduate-employing organizations as a result of those organizations' collective perception that the program is useful in meeting organizational needs. An instructor who is regarded as having conducted a particularly good course may receive a university teaching award. A funding body may decide to support a well-received program and to withdraw support from one which is "unsuccessful". In certain circumstances all of these may be characterized as "outcomes", "effects", or "impacts" of the program itself. They might also be seen as "consequences" which arise because of certain program outputs. The outputs of a program of instruction which are directly related to theories of student learning have been

classified by Doyle as: cognitive, affective, performance skills, personal/social, and vocational outcomes.[16]

Cognitive outputs have to do with the student's imputed ability to perform particularly intellectual tasks. He may, for example, be said to recall facts, understand theories, analyze situations, synthesize action plans, or evaluate instructional programs. Further, to the extent that he is imputed to be able to perform these tasks better as a result of his participation in an instructional system, the system may be regarded as "successful". Psychometrists, curriculum specialists, educational researchers and evaluators have expended huge amounts of time, energy, and money trying to devise means for testing the extent to which such cognitive activities actually do change as a result of instruction.[17] Attempts to measure these sorts of outputs have, in fact, become a major preoccupation and focus for certain types of instructional evaluation efforts.[18]

Affective outputs represent those consequences of an instructional process which are viewed as changes in the student's value and belief systems. As a result of his involvement in an instructional program it is believed that a student may become more open to, appreciative of, interested in, or committed to a particular value orientation; or that he may take on certain beliefs that he did not hold before participating in the instructional process. If such affective changes were intended, and if it can be demonstrated that they actually took place as a result of the program, the program may be characterized as "successful"

as regards that particular constellation of outcomes. As social workers know full well, however, affectivity can be extremely difficult to measure. Educational technologists and psychometrists have diligently applied themselves to some of the problems involved, with the result that certain forms of measurement are available for the evaluator's use.[19] Unfortunately the difficulties involved in conceptualizing and measuring affective changes are extremely persistent and, in relation to the resources that have been expended, headway has been slow. Even the most rigorously validated tests are open to a good deal of criticism. Furthermore, useful standardized tests specific to the needs of social work instruction remain a rarity, and the adaptation of standardized tests developed for other purposes often poses serious transfer problems.[20]

Outputs represented as "performance skills" are particularly relevant to sequences of social work education where the student is ostensibly being trained in preparation for professional practice. We might, for example, wish the student to learn certain skills in verbal interpersonal communication. This is a common sort of objective in both graduate and in-service programs for social workers. If we can demonstrate that the student probably acquired such skills as a result of instruction, then the instructional system may be said to have produced a performance skill output. If the instruction has demonstrably increased the student's ability to write a letter, construct a graph, interview a probationer, influence authorities, or manage

a social service agency, we may say that the increase in ability is a result of his having acquired certain performance skills through the instructional process. It should be obvious, however, that even the most "simple" performance skill is a composite of many different activities and is usually dependent upon the acquisition of certain cognitive and affective potentialities. In terms or its utility to that individual, it may not be necessary that we know the components which combine to produce a particular skill; it is sufficient that he can perform the entire skill activity adequately. However, if we wish to encourage the development of that skill in a student, it may become necessary for us to break the total activity down into simpler parts and assist him in learning each component before he becomes able to synthesize and exhibit the skill in toto. Investigators concerned with the development of business and management skills have considerably advanced our theoretical understanding and ability to teach certain kinds of performance skills (e.g., In-basket Out-basket simulation).[21]

A major difficulty in trying to evaluate a program in terms of its outputs arises from the desire to demonstrate that the "outputs" thought to be consequences of the program actually are causally related to elements within the program. We run up against the complex philosophical problem of "causality".[22] Even if we simply wish to suggest that Y could be caused by X, numerous methodological and statistical difficulties arise in attempting to demonstrate

the nature of that relationship. The testing of hypotheses with empirical data has, therefore, become a preoccupation of many involved in conducting evaluation studies, and the issues which stem from various approaches to hypothesis testing and explanation are at the root of a good deal of the controversy surrounding evaluation methods.[23] It is probable that as our imputed outcomes or impacts become further removed in time and space from the actual instructional program, the difficulties in establishing relationships will become greater. Furthermore, complex outcomes are generally much more difficult to deal with than "simple" ones. For practical reasons, it is one thing to produce evidence strongly suggesting a causal link between an instructional experience and an imputed consequent behavior such as the ability to add 2+2; it is quite another thing to produce strong evidence demonstrating a causal link between participation in an instructional program and the imputed consequent behavior involved in helping a community to organize itself into a more viable social unit.

An approach to evaluation which appears to incorporate the systemic viewpoint is the CIPP model advocated by Stufflebeam and associates.[24] It prescribes evaluation at four levels: context evaluation, which serves "planning decisions"; input evaluation, which serves "structuring decisions"; process evaluation, which serves "implementing decisions"; and product evaluation, which serves "recycling decisions". A thorough summative evaluation, though concentrating heavily on output, would presumably pay some

attention to all four levels. A more formatively-focused evaluation might pay somewhat less attention to outputs and more to processes. It too, however, would necessarily attend to all four levels.

Sources of Information

Whether we are interested in trying to establish causal links, correlations, or useful descriptions of instructional programs, information upon which to base our decisions is needed. It should be a valuable exercise, then, to examine sources from which information might be drawn. Again, we will attempt to do this by concentrating on the people involved in the instructional system or its environment.

If the instructor is a central character in the system, it would seem sensible to view him as a potentially valuable source of information. No one, for example, is better able than he to tell us what his own instructional intentions are; no one is better able to tell us about the practical difficulties he meets in trying to implement a teaching plan; and no one will be in a better position to judge, on a day-by-day basis, the extent to which students are engaging themselves with the subject matter as he visualizes it. One should not assume, however, that the teacher represents the major factor in determining the learning objectives which are possible. He may be able to provide the evaluator with a great deal of information in that regard but, as Bloom, Hastings and Madaus have emphasized,

it is the teaching which determines the objectives which are possible.[25] The instructor may be seen as an instrument by the use of which teaching is accomplished. It is the teacher who initiates the translation between subject structure and learning process structure, and that translation will vary with changes in his approach to teaching. It is within the instructor subsystem that the greatest knowledge of subject structure and learning process structure is likely to reside. It is therefore probable that the instructor will be able to provide relevant information with regard to not only the subject matter but also the instructional materials, teaching methods, and learner activities which comprise the instructional system.

Clearly, the student will be a second major source of valuable information. He too can tell us about intentions--his own, the barriers he meets in trying to realize them, and the dimensions of his engagement in the learning process. Furthermore, since he directly participates in the instructional process, he is in an excellent position to observe and provide information about the instructor, the instructional materials, and other inputs which structure the process. More than any other single source, students are commonly used to provide information for evaluations of instruction. Reviews and bibliographies of student evaluation have been provided by Meehl (1941); Wharry (1952); Morsh and Wilder (1954); Costin, Greenough and Menges (1971); deWolff (1974); and Doyle (1975) to name but a few. Some of the issues involved in using students as data sources

will be discussed later in this paper. For the time being, the reader might note that not only present students, but also former students, can and have been used as extremely potent sources of information about instructional systems.

Administrators, instructor's colleagues, curriculum planners, and instructional specialists also represent potential sources of information. Administrative records, colleagues' observations and judgements, and the theoretical prescriptions and analyses of educational specialists can tell us a great deal about the ways in which a program has been carried out. They can also provide further information about diversely-based intentions for the program. Generally, administration-based and colleague-based information sources have been most heavily used in evaluations which are aimed at producing information for administrative decision-making. The contributions of educational specialists have concentrated more on instructional methods and materials. Neither source of information, however, should be ignored in planning and implementing the evaluation of an instructional program. Administrative and colleague information may cast a totally different light on the instructional sequence. Educational specialists may provide theoretical underpinnings and methodological prescriptions relevant to the examination of a learning process structure.

The employers of students, other instructional agencies, and the clients of students are especially valuable sources of information about the outputs of the system. It is they who use these outputs as inputs to their own systems,

and it is they who are best able to observe and judge the extent to which the products of an instructional system meet their own particular needs. There can, however, be considerable difficulty, both political and methodological, in using these sources of information. To the extent that such organizations and collectives see value in providing feedback to the instructional system, political barriers may be overcome. Practical methodological barriers will still exist and will receive further discussion in later segments of this paper.

Finally, one should not disregard the evaluator himself, or the evaluation team itself, as a potent source of information about dimensions of the instructional system. As we shall see later, some evaluation practices rely heavily on this source of information while others conscientiously ignore it; but, clearly, the evaluator who engages in systematic first-hand observation of an instructional program in operation will gain vast amounts of information; information which is filtered through his own perceptual screen. His own pool of first-hand knowledge may be the best available information source about system processes. An emphasis on the evaluator's use of his own first-hand knowledge, as one of many information sources, is a principle of one of the evaluation paradigms which will be examined in this dissertation.

We may conclude that there are numerous and diverse sources of information that may assist in evaluating an instructional program and informing the decision-maker. To

a large extent, the sources of information which are tapped by an evaluation study depend upon the overtly acknowledged purposes and foci of the evaluator and the decision-maker whom he serves. We will see, however, that in certain approaches to evaluation limitations are placed on the use of some sources, while other approaches emphasize the use of as many sources as is practicable.

Media of Information Gathering

There are many ways of gathering information from these sources and these media will, to a large extent, dictate the sorts of information gathered. Tests, rating scales, interviews and conversations, direct observation, surveys, and analysis of documents represent the major media of data collection. Each must be considered in terms of the range and types of data it tends to produce, as well as the quality of that information. Further, though there are problems common to all forms of data collection, each medium imposes particular problems. I will attempt to outline the major dimensions of these media and examine some of the problems and sorts of information associated with each one. By and large, the question of "information quality" will be left to examination at the end of this Chapter.

There are hundreds of different kinds of tests which can be used for assessing elements within, overall dimensions of, and outcomes of instructional systems. A paper of this sort could not begin to discuss them thoroughly. The development and application of tests, as well as the analysis of

test-produced data, has undoubtedly been a major focus of educational research. In terms of evaluation studies, tests are used primarily in the assessment of inputs and outputs. If, for example, we wish to know about the personality characteristics which a student brings to the instructional system, we will probably devise tests to shed light on the dimensions of those characteristics. We may use those tests againto measure the same personality characteristics when the student leaves the instructional system, and, if we go about it correctly, we may be able to infer whether or not the instructional system has produced any changes in those characteristics. The same sorts of media are often used in the assessment of cognitive variables, affective variables, and performance skill variables--whether they are represented as systemic inputs or outputs.

Literature on the construction, application, and interpretation of cognitive tests is rich and readily available.[26] Although discussion of various issues and problems must concern us later in this dissertation, a thorough review here would be an impossible task. There are existing standardized tests which may prove amenable to the evaluator's uses and it is entirely possible for him to construct his own specifically useful instruments, but testing cognitive variables is an exceedingly complex enterprise which should not be undertaken naively.[27] If a test is to contribute high quality information to decision-making it must be chosen or developed very carefully. A good test is the product of years of development--years in which its usefulness

grows as experience with it increases.

Much the same may be said about tests designed to assess affective variables; although less success might be expected in this probem-ridden area. The literature, however, is broad and recent concise reviews can be found in Travers (1973, see Ch. 24) and Bloom, Hastings and Madaus (1971, see Ch. 10).[28]

The measurement of performance skill variables does not appear to have received so much attention in the educational literature; perhaps due a belief that skill variables are not so necessarily ingrained in educational, as opposed to training, pursuits. However, Simpson (1966) has attempted to extend Bloom's Taxonomies by consideration of psychomotor objectives. The increased emphasis on behavioral definition of learning objectives has added a great deal to this literature in the past ten or fifteen years.[29] For practical guidance in the testing of performance skills, probably the best source is the quickly-growing literature on industrial and management training.[30]

Whatever the subject matter of tests, they must be selected or developed with regard to the specific purposes of an evaluation, and, whenever possible, they should be tried out in the particular situation where we intend to use them. As Taylor notes, there have been too many occasions in which testing programs have been instigated without tryouts and the results have led to unsound decisions.[31] Unfortunately, particularly in small-scale and "one-off" programs, it is often impossible or impractical to adequately

try out a test or to compare the utility of several tests in preliminary studies. Finally, it must always be remembered that tests are merely instruments for producing information; they will not make decisions for us.

As Doyle points out, ratings are essentially multiple choice questions for which the response options describe judgements or observations--generally about instructor traits and behavior, material, outcomes, or environmental variables.[32] Student rating scales have enjoyed wide application in the evaluation of instructional programs, and their frequent use stems from the fact that they are economical, speedy, relatively simple to design, and readily summarized and recordable. They do, however suffer from a number of drawbacks and should always be supplemented by other sorts of data.[33] Many rating scale items are of the general or "high inference" type (e.g., "Please rate the instructor's over-all teaching ability"), which may be excellent in terms of their technical properties but offer very little information useful in attempting to improve instructional programs.[34] Guilford (1954), Berdie and Anderson (1974), and Doyle (1975) provide thorough discussion of issues and methods in the development of rating scales.[35]

Of course, students are not the only possible sources of rating data; instructors, administrators, and colleagues may also provide ratings. A number of studies have attempted to compare ratings from these various sources. For example, student ratings of instructors have been compared with instructors' self-ratings by Webb and Nolan (1955), Blackburn

and Clark (1971), and Centra (1972). On the whole, these studies suggest that, while such ratings sometimes agree with each other, and self-ratings are sometimes more severe than student ratings, instructors tend to rate themselves more favorably than do their students. Comparisons of student, colleague, and administrator ratings suggest that all three tend to agree about the general characteristics of instructors but tend to disagree on more detailed classroom-specific characteristics.[36] Two studies of differences between present student and alumi ratings conclude that they differ very little.[37] When there are differences, it is the alumni who tend to give less favorable ratings--perhaps supporting the presumption that instructors tend to improve with experience. Doyle (1975) notes that there is similarity between rating data from all of these sources but suggests that it is not strong enough to allow for the substitution of one source with another.[38]

There has also been some study of the relationship between student ratings of instructors and student characteristics; particularly demographic, ability, and motivation characteristics.[39] Neither students' demographic characteristics nor their ability characteristics show any systematic relationships with student ratings of their instructors. The motivation-related studies found that ratings of instructors were generally less favorable from students for whom the course was "required" rather than "elective", but this distinction does not constitute a clear-cut

separation of students according to their motivations; an instructor's behavior may differ between "required" and "elective" courses. Thus, generalizations about relationships between student motivation and student rating of instructors should be regarded with a good deal of suspicion.

It is widely believed that raters will rate differently if they are identifiable, rather than anonymous. Sharon (1970), however, found that although identifiable students may rate instructors slightly more favorably, the differences in average ratings were not statistically significant. Anonymity is usually prescribed, however, because it is believed that raters will probably be more willing to provide data if they are not identifiable.

Finally, at least two studies have found that rating data seem to be significantly influenced by the believed purpose of the evaluation.[40] For example, when students believed that their ratings of instructors would be used for personnel purposes, their ratings were more favorable than when they believed that the data were for research purposes only. Since the believed purpose of the evaluation does seem to affect ratings, evaluators would be wise to ensure that they do not represent the purposes of their study differently to various sub-sets within the population of raters.

In conclusion, although there are many controversial issues and difficulties associated with rating scales, much of the literature suggests that they may prove to be a very useful medium for obtaining information relevant to

the evaluation of instructors, and probably other elements in the instructional system as well.

Although disregarded entirely in some evaluation designs, interviews and conversations with participants in the instructional program can provide volumes of rich, illustrative information which may prove to be of inestimable value to the decision maker who must develop an integrated understanding of the system. These media can be used to produce information which is more clear and complete than that resulting from the use of any other medium of investigation. Some of their obvious uses would be in: exploring basic values and attitudes held by participants, exploring changes in expectations and life plans which occur in the student, and gathering participant impressions about characteristics of the system. Undoubtedly the information gathered with these flexible media constitute both a broad and deep resource for answering questions which occur before, during, and after program implementation; however, the data produced is often highly unweildy. As we shall see later, the task of organizing culling, and making sense of this sort of data may be an awesome one, and any evaluator intending to use these media would be wide to develop some data selection criteria before applying them. There are many points of controversy surrounding the use of interviews and conversations as data collecting devices; some of them will emerge at later points in this paper.

A variety of methods for the direct observation of instructional programs in action have been developed, mostly in connection with research on student teaching and primary school instruction.[41] Many evaluations tend to ignore them. Critics argue that such procedures are expensive and require special skills in the observer; that they constitute an invasion of privacy resisted by instructors; and that the observer's presence strongly influences participant behavior in the instructional process--to the detriment of the instruction itself and the evaluation. Proponents of direct observation admit that it can be a difficult and expensive exercise, and that it may create additional generalizability problems, but believe that neither of these criticisms invalidates its use. They argue that once instructors are committed to evaluation for program improvement, their resistances will weaken. As for the third criticism, it has no completely satisfactory answer, but as Medley and Mitzel note, "To know how teachers and pupils behave while they are under observation seems better than to know nothing at all about how teachers and pupils behave".[42] Because other media do not provide particularly satisfactory means for identifying specific patterns of effective instructor behavior, the systematic direct observation of classroom behavior has become part and parcel of "methods research" in education.[43] It could well eventuate that adaptation and application of the more complex and "sophisticated" observational systems will be too expensive for use in many evaluation studies. However,

this should not prevent us from developing and using more simple models.

Almost every instructional system will have existing associated documents, such as written syllabi, program proposals, correspondence between participants, examination scripts, time-tables, students' notes, and consultants' reports which may be examined and analyzed to provide information useful in evaluation. Furthermore, certain kinds of documents, like student essays, participant diaries, and instructors' notes, may be encouraged and produced especially for evaluation purposes. Gathering and analyzing this sort of information may provide historical perspective, indicate areas for inquiry, and expose aspects of the program which might otherwise be missed.[44]

Survey questionnaires are most often put to use in educational evaluations to gather information on the personal and social characteristics of students, or in broadly gauging student reaction to elements of the instructional system either during or after their involvement in the process. Sometimes they are used in lieu of interviews because they are less time consuming, less demanding of administration skills, and less expensive. In some instances they are preferable to interviews because they pose questions uniformly to all subjects and they can allow for subject anonymity. However, they cannot provide the "depth" of information garnered in a good interview, they require a relatively high standard of reading and writing skills on the part of the subject, and they are not particularly

good for gauging the strength of beliefs and attitudes. There has been a great deal of publication in the area of survey and questionnaire design.[45]

Quality of Information

We have assumed that good decision-making requires good information and have briefly reviewed some of the sources and media available for the production of information about instructional systems. We are now led to the question, "What are the characteristics of good information?" The literature points to at least four major characteristics that should be examined: validity, reliability, generalizability, and utility. Whatever its source or media of generation, information for decision-making should be evaluated with regard to these characteristics. Furthermore, it may be argued that certain kinds of program decisions (e.g., personnel decisions) require a higher or different quality of data than do others (e.g., program improvement decisions).[46]

Definitions of validity differ but, generally, information is considered to be valid to the extent that it accurately represents its subject matter.[47] For example, data derived from a questionnaire item on employment experience are valid to the extent that they accurately represent that employment experience. If we should find that such a questionnaire item tells us, perhaps, only what the subjects want us to know about their employment experience, then the item would not be considered very valid.

Of course, it is not quite that simple; validity may be viewed as existing on a continuum. That is, instruments or information are not necessarily either valid or invalid, one may be judged to be more valid than another. There are three commonly used sorts of validation: content validation, empirical validation, and construct validation.

"Content validity" is reflected in the degree to which a test is a representative sampling of a segment of the behavioral domain we wish to assess.[48] To illustrate, in evaluating an instructional program on automobile production, we may develop an instrument which will be administered to assess the students' acquisition of knowledge about automobile manufacturers. That instrument would be considered valid to the extent that it tested the knowledge areas covered by the program. If the content of the instructional sequence on automobile manufacturers was 60 percent taken up with Ford, 30 percent with Chrysler, and 10 percent with General Motors, then the requirements of content validity would expect the instrument to reflect this 60:30:10 ratio. To the extent that the ratio was not reflected in the instrument, that instrument would lack content validity on that dimension. If the ratio reflected was 55:35:10, then the instrument might be judged to have relatively high content validity on that dimension. If the reflected ratio was 20:10:70, then the instrument might be judged to have relatively low content validity on that dimension. If the test dealt with motorcycle manufacturers rather than automobile manufacturers, it might be judged to

have absolutely no content validity. Obviously an instrument which set out to test a range of knowledge from a much more complex instructional program would require assessment of content validity on many dimensions, but the principle remains the same. Content validity depends upon the extent to which the dimensions of the instrument, or information, reflect the dimensions of the instructional program elements under study. Chase (1974) outlines a grid method for applying Bloom's taxonomy of cognitive objectives to planning or evaluating a unit of instruction. The use of procedures such as this could greatly assist in judging the content validity of our evaluation media and information. Another procedure which could be used for the same purpose is factor analysis. Briefly, it is a mathematical procedure for discovering the underlying themes or dimensions of a set of variables. Doyle has summarized a number of studies which used factor analysis to examine the content validity of student ratings of instructors, and discusses the innate subjectivity of content validations.[49]

Instructional programs are very often aimed at the development in students of characteristics such as "mathematical ability", "creativity", "communication skill" or ability to practice". These characteristics are referred as "constructs". When we identify clumps of behaviors which seem to be related to each other, we often attempt to develop a theory about an underlying trait or "construct" which cannot be directly perceived but seems to organize, control, or be the source of the apparently related groups of

behaviors. In order to evaluate an instructional system we may wish to assess evidence with regard to the validity of such constructs. Chronbach and Meehl have suggested sources of evidence about construct validity which provide guidelines for the evaluator who wishes to take measurement of such unitary traits.[50]

We noted the innate "subjectivity" of content validation. Empiricial validation attempts to use more "objective" measures as criteria for assessing the validity of information and the media we use to produce it. If, for example, our program is aimed at helping the student to "perform all the cognitive tasks expected of a practicing professional social worker", this "ultimate" criterion will prove to be too difficult to measure with any degree of objectivity. In such cases we can substitute "intermediate" criteria such as test score rankings, measurement of physical responses, observed performance of specific tasks, and the like, which are reasonably related to the "ultimate" criterion. We may not be able to measure "student learning" as a whole but we may be able to develop and use relatively objective instruments which strongly suggest the degree to which a particular element of student learning is present or absent. Statistical correlations are widely used for establishing the degree of relationship between "objective tests" and ultimate or intermediate criteria.[51] It should be noted, however, that even this highly quantitative device is often of questionable value because we seldom have good reasons to expect a linear

relationship between test scores or ratings and student learning. For this reason, it has been suggested that wider use be made of curvilinear statistics and "threshold measures".

Finally, in all validation, one should not assume that some procedures are "subjective" and others "objective". Procedures are subjective to the extent that they rely on judgements or reasoning on the part of people; they are objective to the extent that they use sound empirically-determined facts and, often, quantitative techniques. They must, however, be regarded as existing on a continuum between "absolute objectivity" and "complete subjectivity". All validation, as all evaluation, includes both fact and judgement. Discussion of this continuum will necessarily arise later in this paper.

The concept of reliability in data and data-producing instruments has been dealt with extensively by a number of writers and will receive only a general review here.[52] It may be generally characterized as "the extent to which information is consistently accurate". Obviously, it cannot be automatically assumed that any instrument we use will provide consistently accurate information. For example, a test might provide a measurement which is consistently too high or too low; these are called "constant" or "systematic" errors. In the same way, a particular group of students might consistently rate an instructor higher or lower than his customary performance would entitle him to. In these examples, we would question the accuracy of the

information produced. There are also occasions when the information is not consistently accurate. For example, a student's emotional reaction to the circumstances surrounding the application of a test may influence his performance, or a particular student may rate an instructor very low on "teaching ability" when it is really the teacher's daughter he dislikes. These are called "random" or "chance" errors because they are essentially unpredictable and they are not displayed by all subjects in the same way or to the same degree. The standard assessments of random error involve examining internal consistency (homogeneity) and retest reliability (stability).

If it is our intention to measure the same "thing" with various questionnaire items, test questions, or ratings, then a high degree of internal consistency between items, questions, or judgements suggests that we are receiving consistent information. If we use those measures over and over again on essentially the same subjects and find that they give us very similar information each time, then we have another indication that the information is consistent. These "internal" and "external" consistencies are independent of one another. A number of procedures for the statistical computation of random errors are reviewed by Thorndike (1967), and Chase (1974).[53]

High assessments of neither internal nor external consistency are enough to convince us that <u>systematic</u> or <u>constant</u> errors are not occurring. Since systematic errors are consistently inaccurate, they are extremely difficult

to assess. Doyle reviews a number of sources of error in rating data, many of which are systematic.[54] Since no general method is available for assessing systematic error, investigators have concentrated on identifying various types and developing methods particular to the assessment of each. Guilford has proposed an approach to estimating and correcting for "leniency" and "halo" effects; work on "proximity errors", "logical errors", "contrast errors", "sharpening", and the effect of expectations have been dealt with elsewhere.[55]

The reader will remember from our discussion on "foci" of evaluation that many evaluation studies tend to concentrate on gathering information only about rather specific inputs (such as the instructor's behavior) or outputs (such as students' cognitive achievement). Such limited foci impose limitations on the information to be gathered and consider characteristics of a very limited universe of information for which the evaluator has special interests. When we attempt to gauge the extent of certain cognitive achievements, for example, we take a sample from the very large universe of information about cognitive achievements which is potentially available. That sample of information would constitute a very small segment of all potential information about the outputs of that particular instructional program. Further, the sample would constitute an infinitesimal part of all potential information about the outputs of instructional programs in social work. Clearly, evaluation data are always a sample from some universe of

data and, like other samples, their goodness may depend upon the accuracy with which they represent the universe of data in which we are interested. To illustrate, if we are interested in knowing about all the effects of an instructional program concerned with teaching social work principles, a study which estimates the extent to which students learned to understand the principle of confidentiality would provide only a very small sample of the universe of information which interested us. If, however, our interest was only in information regarding the effects which the program had on students' understanding of casework principles, the same sample would be a better representation of our universe of interesting information. If we were specicially interested in the effects which the instruction had on students with regard to understanding the principle of confidentiality, the same sample might be an excellent representation of that universe. Likewise, if the taking of certain decisions required that we have as much information as possible about all aspects of an instructional program, then information about program outputs, processes, or inputs alone would poorly represent the entire universe of information which was required. Before undertaking an evaluation, then, it is wise to develop some guidelines with regard to the sorts of information decision-makers need or desire. Do decision-makers want information about program outputs which are related to a particular closely-specified objective? Do they want information about outputs relating to a number of program

objectives. Perhaps they most need information about certain instructional processes or inputs. We must, to some extent, define our universe of useful information before we start developing and applying instruments; otherwise we are likely to end up with a sample of information that will not permit us to make the sorts of inferences and generalizations we wish to make. The goodness of an information sample may be judged against the same sorts of criteria that we use for other types of samples, and the literature on sampling is enormous.

In summary, any information gathered about an instructional program, by whatever means and from whatever sources, is a sample of a universe of possible information about that program. This means that any piece of information is of limited value. It should be clear that we can validly generalize from our information only to the extent that our sample validly represents the universe to which we wish to generalize. Generalizations about the effects or overall effectiveness of a program from information about the effects or effectiveness of a minority of program elements may, therefore, be extremely misleading. By the same token, generalizations about the effectiveness of all instructional programs in social work from information about the effectiveness of a particular instructional program in social work may be absurd. Estimations of the extent to which such generalizations can be valid and useful should be based on established principles of sampling theory.[56] In passing, it might also be noted that samples of information are

sometimes legitimately drawn with complete disregard for statistical representativeness, usually to explore some special issues or concerns. We will return to this point in another Chapter.

As Weiss notes, "Evaluation as an applied research is committed to the principle of utility. If it is not going to have any effect on decisions, it is an exercise in futility."[57] Therefore, the information which is produced by evaluative investigations must be regarded as useful to the extent that it serves decision-making purposes. Information-gathering devices may be regarded as useful to the extent that they provide useful information. Their utility may be assessed against a variety of criteria which, of course, also relate to our purposes. It is one thing to develop information-gathering instruments that, theoretically, will provide us with valid, reliable, and generalizable information. It is quite another thing to develop instruments that will actually prove useful to decision-makers. In their preoccupation with validity and reliability issues, some "evaluators" seriously neglect examination of this dimension. Those who wish to perform evaluation studies must consider the utility of their proposed instruments. Such considerations appear to be largely commonsensical; but when one begins to more deeply examine criteria such as economy, ease of administration, and ease of interpretation, the examination can become very complex.

If the information obtained through a medium is not worth the expense of acquiring it, then that medium is not

of much use to us. In this respect, "expense" can be defined and measured in many ways, depending upon the sorts of resources that are necessary and available. An instrument which is economical in terms of its drain on financial resources might be entirely too expensive in terms of the time and effort used in administering it. Furthermore, ease of administration may be governed by a wide variety of considerations other than money, time, and effort. For example: is special knowledge or skill necessary in applying the instrument; will subjects cooperate to the extent necessary; does administration require further research before it can begin to operate? In addition, should we find that we have an instrument that can be administered relatively efficiently, we may find that the information produced is too unweildy for economical analysis and interpretation; alternatively, that there are ethical considerations which mitigate against its use. Although proposed information-producing techniques may be studied theoretically and possible difficulties predicted ahead of time, it is only through actually using them that some of these contraindications become apparent.

If our purpose in evaluation is the "improvement of instruction", a certain range of information will be useful to us. If we define our purpose as "promoting student learning", a somewhat different range of information will be useful to us (i.e., "student learning" may not be the only objective of instruction). Some studies have been concerned with the utility of student rating evaluations

in promoting instructor improvement.[58] They generally indicate that evaluations using student ratings do influence instructors toward improvement (as defined by students) so long as the evaluated qualities are changeable, the instructor is motivated to change, and the evaluations provide explicit guidance regarding the changes desired. Note, these studies do not attempt to tell us anything about student learning; they merely suggest that student rating evaluations can influence instructor behavior toward improvement as defined by students. In order to establish a relationship between "improvements in instructor behavior as defined by students" and "improvements in student learning" we must seek direct evidence of it in other ways. Even if we are satisfied with information suggesting an influence toward instructor improvement, we may conclude that rating scales are generally only somewhat useful because their breadth often operates against useful specificity. Doyle suggests that narrative evaluation media (e.g., conversations and interviews) and good outcome measures (e.g., well constructed cognitive tests) would probably provide an instructor with more specifically useful information upon which to make decisions about changing his own teaching behavior.[59] The utility of student ratings and other media of data generation will, of necessity, be discussed further in later sections of this paper.

This chapter has attempted to provide a broad overview of the evaluation of instructional programs by examining (1) a limited number of definitional elements, distinctions,

and issues; (2) some of the purposes and foci which are often associated with evalutions; (3) information sources and media of generation; and (4) a few considerations with regard to instrument and information quality. It has neglected an array of important issues, some of which will be examined in later Chapters, but provides a base upon which further discussion can be developed.

CHAPTER II

THE LITERATURE ON EVALUATING SOCIAL WORK INSTRUCTION

Any review of the social work literature which attempted to deal only with material exclusively aimed at the evaluation of instructional sequences in social work would be very limited indeed. Although a very thorough search would undoubtedly yield more evaluation studies than are reported here, such a survey would necessarily contain a great deal of near-duplication. Furthermore, the majority of reports, articles and volumes which appear to offer information valuable to the task at hand do not concern themselves specifically with evaluation as such. Curriculum design proposals, theoretical formulations of social work-linked knowledge and values, treatises on the components of social work practice skill, discussion of teaching methods, examinations of teacher and student role expectations and performance, and a host of other enquiries and expositions on tangentially related subjects combine to form a background against which approaches to the systematic evaluation of instructional programs in social work have slowly developed. Since most evaluative studies address themselves to a relatively limited audience, many are never published. Further, it may be that social work's traditional emphasis on a "caring focus" and social work practice as an "art" have tended to reinforce a predisposition against measurement and scientific evaluative research. In any event, exclusive examination of the few evaluative

studies which do exist would appear to be wholly inadequate. The intention in this Chapter, then, is two-fold: to chart some of the highlights--some of the contributions which have led toward systematic evaluations of social work instruction; and to examine a few examples of the evaluations that have been published--as well as a smattering of the related research literature. It is hoped that the overall effect will be to acquaint the reader with a broad overview of "the state of the art".

We will begin over thirty-five years ago with Bertha Caplan Reynolds who developed an initial analysis of the functions of the social work teacher.[1] Although social work practice and social work education have undoubtedly undergone considerable changes in the intervening years, much of her discussion is relevant to the activities of today's social work instructor. Her examination was, in fact, a beginning attempt to formulate principles which would guide the teaching of social work practice. She discussed the composition of educational "diagnoses", means of presenting subject matter, providing the learner with security so as to encourage exploration, the encouragement of realistic and direct study of problematic situations, and the concentration of instructor effort on the learner. At that early stage in the development of modern social work the systematic evaluation of instructional activities appeared to be a long way off. However, Reynolds' discussions demonstrate that there was a great deal of concern for the development of instructional modalities which would

lead to the eventual improvement of social work education and, ultimately, to improved social work service.

In 1950 Helen Harris Perlman presented a rather thorough examination of the "discussion method" in teaching casework.[2] In the following year she published a similar discussion on the "lecture method".[3] These articles contributed substantially to a growing concern about examinations of instructional procedures, and it provided exceptionally well-detailed conceptualizations of conventional practices in social work education. The contention that these two articles signal a growing concern for examination of teaching methods is supported by the concomitant appearance of several similar articles exploring approaches to instruction and culminating in the "Curriculum Study" carried out by the American Council on Social Work Education.[4]

In 1952 Ralph Tyler, a leading educational researcher, addressed the Annual Meeting of the American Association of Schools of Social Work.[5] His overview of the essential characteristics of a profession from an educationalist's standpoint identified some of the major tasks involved in planning and conducting an educational program. No thorough summary will be undertaken here. What is of primary interest is that one of his identified central "tasks" was seen to be "evaluating the effectiveness of the educational program in attaining its objectives through appraising the educational progress of the students." Completion of this task, he suggested, required student assessments early in the

the course as well as near the end, and involved the gathering of evidence relating to all of the program's important objectives in an effort to identify those aspects of the curriculum that are effective and those that require improvements. He concluded his presentation with an examination of four important professional attributes, indicating subtasks: (1) appraisal of all important objectives to develop a comprehensive picture of the achievements of students in relation to the school's purposes; (2) the use of varied devices for obtaining evidence regarding students' educational progress; (3) evaluation at several points in the student's career--minimally, one early in his training period, one near the end, and one after several years of professional service; and (4) conscious use of evaluation in improving the educational program. This would appear to be the first discussion to appear in the professional literature which presented a detailed approach to the evaluation of instructional activity.

Soon thereafter, Charlotte Towle published The Learner in Education for the Professions in which she made explicit use of "theories of learning and teaching" as she developed a set of principles to guide instructional activity in social work.[6] Her examination provided a seminal analysis of the role of the social work instructor while, at the same time, concentrating on the needs of the learner. In the process she made note of a serious need for research to guide instructional development. "It is evident", she said, "that to understand the learning process of adult learners

in a profession, we must study their responses to the professional education situation. Professional education must do its own research. The findings of social work educators at present are equivalent to those of the experienced clinician, unsupported by research." Further, she went on to presage a very serious difficulty in the conduct of such research. ". . . the situation cannot be isolated from the life situation. . . . Hence, there are limits to the controls that can be set in the research situation. We have the problem of studying change in a dynamic situation."[7]

Unfortunately, it would appear that no great research boom followed from the calls of Tyler and Towle. As mentioned earlier, quite a number of exploratory and formulative articles followed, but research studies enquiring directly into instructional practices remained a rarity. One of the few articles which did make a substantial research contribution was an enquiry undertaken by Alfred Kadushin and reported in 1956.[8] Interestingly, it did not follow from the approach suggested by Tyler. Kadushin was interested in student observation of interviews as a teaching device, and appears to have realized that exploratory work would have to be done before an experimental study could be considered. Accordingly, he used a questionnaire to determine the relevant opinions of a sample of social work faculty members, public welfare in-service training consultants, and executive directors or casework directors of large family service agencies. He then summarized the data collected and used numerous direct quotations from

the responses to describe and explore the various issues and viewpoints which had been expressed. The result was a relatively comprehensive, richly illustrated, and very interesting examination. It indicated the extent to which direct observation of casework interviews was being used as a teaching resource and explored the major affirmative and negative arguments from a number of viewpoints.

In 1958 Martha Moscrop, in her well-received volume on in-service training, devoted a chapter and an appendix section to "evaluation".[9] The emphasis, however, was firmly on "evaluation" as assessment of the social work practitioner--aimed at improvement of the practitioner's work. Although she noted various administrative considerations and staff development implications relating to the assessment of practitioner behavior, she surprisingly neglected to discuss the utility of such activity in providing a baseline for curriculum development and the evaluation of instructional programs.

In a similarly tangential, but equally valuable, piece of work Elizabeth Herzog entered into a very detailed discussion of evaluative research in social work.[10] Again, the discussion was not specifically aimed at identifying and explicating issues and problems in the evaluation of social work instruction. It did, however, deal with the evaluation of efforts to bring about "psycho-social" change in individuals and, therefore, contained numerous apposite comments and suggestions. The nine questions used to organize her discussion appear to be as relevant to the

evaluation of instructional systems as to "treatment" systems.

(1) What is the purpose of the evaluation?

(2) What kind of change is desired?

(3) By what means is change to be brought about?

(4) How trustworthy are the categories and measures employed?

(5) At what point is change to be measured?

(6) How fairly do the individuals studied represent the groups discussed?

(7) What is the evidence that the changes observed are due to the means employed?

(8) What is the meaning of the changes found?

(9) Were there unexpected consequences?[11]

By ending with some very specifically detailed guidelines, Herzog provided not only an excellent analysis of evaluation issues but also an extremely useful prescriptive synthesis. She also emphasized that fantasies of "neat, precise, utterly objective social science modeled after a naive conception of the natural sciences" were unrealistic, and that there was value to be had from seeking out skilled expert judgement and frank opinion rather than "hard proof".[12] Her work established a richly detailed substructure upon which further evaluation efforts could be developed. The bulk of the studies which did follow concentrated upon evaluating the effectiveness of direct services to social work clients; but, as Sanders notes: "Any consideration of the preparation of social workers to be effective change agents necessitates

focus on methods for the evaluation of the effectiveness of social work intervention."[13] Unfortunately, a review of the growing literature in that area would be entirely impracticable here.[14]

In 1964 Marguerite Pohek spurred the social work education community with a general examination of teaching methodologies which related to the questions: "To what end are we educating?", "Whom are we educating?", and "How are we educating?". Her examination appeared, in part, to be a flowering of the seeds planted by Towle a decade previously, in that it clearly supported andragogical rather than pedagogical principles and concluded with some very specific recommendations toward a more concerted involvement in auditing and assessing instructional programs.[15] Over the years social work education had increasingly become viewed as a distinct specialism within the broad sphere of social work activity. However, the desired research accompaniment did not appear to have developed to any significant degree. In her discussion of "routes" for the preparation of faculty which appeared in the inagural issue of the Journal of Education for Social Work Blackey laments this deficiency.

> Actually, with the introduction of social work education as a specific area of specialization, research within the specialization could be focused on the design, testing, and evaluation of approaches to learning theory, curricula development, and teaching methods. One reason for our limited advances in social work education as a field is our lack of research into these processes, which currently receive attention in our doctoral programs to a negligible extent, if at all.16

The new Journal of Education for Social Work provided social work educationalists with their own forum for

specifically relevant discussions and investigations. That first issue, for example, contained a number of other articles which pleaded for more evaluative research and contributed to the impetus toward investigations which were specifically relevant and useful to social work instructors. Alfred Kadushin contributed a discussion of objectives for social work education which emphasized the concern that social work students should learn to do something rather than simply concentrate on developing "disciplined self-awareness".[17] Herman Stein pointed out the differences between graduate-level education in social work and in-service training.[18] While encouraging schools of social work to assist social service agencies in conducting more effective staff development programs, he strongly advocated that the schools should effectively utilize their scarce resources in developing broadly preparatory professional education curricula rather than programs aimed at introducing students to particular tasks in particular agencies. Mary Burns briefly reviewed the state of social work knowledge and revealed some of the pressing knowledge needs of the profession.[19] The consequences for curriculum-building and evaluation are readily apparent in one of her conclusions:

> A review of the current state of social work knowledge discloses a devastating lack of agreement as to the definitions of many of the concepts in use. Frames of reference are vague and general, and theories are either so comprehensive that they are difficult to render into practice principles or so limited at the practice principle level that they are seldom applicable. Middle-level theory, from which practice principles can be derived, is largely missing. Practice knowledge has been only inadequately codified, and practice wisdom is largely

> unvalidated. The net result is that, in many of the practice methods by which social work seeks to achieve its ultimate aims, professional social workers must function with a body of knowledge that at best is imprecise, at worst is erroneous, and always is inadequate and difficult to communicate.20

Undoubtedly, the article in that volume which was most closely related to our investigation was provided by Samuel Finestone.[21] Finestone presented a perspective for research on social work education based on the exploration of two fundamental questions: "What is an appropriate way of viewing social work education as a whole?" and "What research questions are generated by this comprehensive overview, and what classification scheme usefully orders them?". The resulting model of social work education might be seen as little more than a relatively tidy academic exercise were it not for Finestone's detailed classification scheme for research questions. He eventually broke down sections of the model to formulations of research areas which suggest rather specific directions for investigation. For instance, under the heading "The Graduate School as a System of Educational Processing" he suggested a category of research labelled "Descriptive Studies of Educational Structure, Program and Process". He then went on to flesh out the category, stating that it ". . . includes comparative studies of objectives, content, methods and organization for the various course concentrations and fieldwork programs in schools of social work."[22] Further, he outlined some of the possible benefits which might accrue from such research. Another category in the formulation was lablelled "Analytical

Studies of Structure, Program and Process", and this category was further broken down; one component being:

> . . . analytical studies focused on evaluation of effectiveness. . . . In such studies, desired student performance is the major dependent variable (performance viewed from many points of view); and the questions revolve around the contribution of educational program aspects. For example, segments of program (particular methods, or experimental programs, or organization and spacing of curriculum) may be the basis of assessment. The methodology of educational evaluation, as well as evaluation of program aspects and total program, would seem to deserve attention.23

Categories of research activity which logically fall under other major headings are considered as well. Thus, by beginning with two basic questions, formulating a comprehensive model of social work education, and inexorably honing down categories of research questions, Finestone develops a wide-ranging, yet very detailed, program for proposed social work education research which suggests, among other things, a variety of approaches to evaluative investigation. Carter and Wharf have undertaken a more recent and somewhat similar exploration which leads toward a taxonomy of evaluation methods related to social welfare programs.[24]

Working on a rather different and more circumscribed front, Margaret Heyman attempted to formulate a set of "criteria and guidelines" for the evaluation of in-service training for social welfare personnel.[25] Her initial clarification of definitions and basic assumptions was a useful contribution in itself, but her discussion of criteria and guidelines related to adequacy, content, methods, and effectiveness combined with her elucidation of general

approaches to evaluation to form an excellent exposition of the issues, pitfalls and problems inherent in many efforts to evaluate social work instruction. She noted that social work ". . . lacks a simple and useful system of case classification tied in with specific techniques of intervention that can be used as a base to measure client change", and that ". . . measurements of organizational change are still rudimentary". Therefore, she suggested that it was necessary to examine intermediate objectives of training. Clearly, Heyman's volume makes many useful suggestions and explores a variety of educational issues in relation to social work content; however, it has one serious shortcoming. Her entire discussion remains solidly within the traditionally-espoused "scientific" experimental approach to investigation. By neglecting to consider possibly fruitful alternative approaches to evaluation, it remains a rather limited enquiry.

In the same year Dorothy Pettes concentrated on the supervision process, and clearly outlined three conventionally recognized aspects involved in learning social work practice: the accumulation of factual knowledge, the learning or modification of attitudes, and the development of professional skills.[26] She also took note of the social worker's need for continued education throughout his professional career and expostulated on some "basic principles" of education. However, she did not develop these principles into specific criteria for use in assessing student performance or instructional effectiveness.

In 1969 Mary Louise Somers extended Reynolds' and Towle's formulations on the role of the social work teacher and examined common assumptions about learning and teaching, as well as some of the current theories of learning as they applied to learning social work practice.[27] She reviewed relevant principles derived from Stimulus-Response Theory, Cognitive Theory, Motivation and Personality Theory, and Gagné's Learning Theory. Somers was, of course, unable to identify one clear and comprehensive theory of learning. In practice, she said, we cannot wait for such an encompassing construct. It is necessary to select the plausible and helpful, and use it inventively. "Thus teachers test fragmentary or partial theories of learning and teaching in the crucible of teaching practice, and, at times, contribute in this way to the advance of educational theory."[28] She ends with the now familiar admonition that more investigative activity must be vested in the learning-teaching transaction itself.

Although far from a detailed history, our review so far indicates the status of evaluative research into social work education at the end of the last decade. There had been a modicum of basic exploration with regard to instructors' roles, learning theory applications, and overall curriculum development needs, but very little actual research had been published. The studies which had been undertaken, principally in schools of social work, were relatively localized and no useful publication of results had emerged.

In 1970 the Council of Social Work Education published Teaching and Learning in Social Work Education, compiled by Marguerite Pohek.[29] As a whole, the compilation provided a refreshing look at social work education from a variety of relatively new perspectives. N. L. Gage's contribution presented a resumé of empirically-based research findings on teacher behavior, with stress on "successful" teacher behavior.[30] He noted some reasons for pessimism about science's contributions to the art of teaching, then presented evidence which questioned that pessimism. Eventually, he challenged the profession with the following statement.

> Without research on teaching and learning, social work education must forever debate its issues on the basis of hunches and intuitions. If such methods of formulating the ways in which social workers get educated have proven adequate and satisfying, they should continue. But if you feel that hunches and intuitions could benefit by the use of hard evidence gained through research, then you will join in the educational research movement. From my own point of view, as an educational research worker, such a decision by your profession would be all to the good.31

In the same volume Florence Clemenger raised some basic issues and questions concerning the assessment of student performance, and attempted to explore some possible solutions. She discussed "objectives", "methods of assessing student performance", and "grading". Finally, she concluded, "Our lack of background knowledge in educational test construction and evaluation techniques has not been resolved even though we have an increase in the number of doctorates on our teaching faculties. . . . This is an area to which some attention is currently being given, but

more needs to be done if we are to know not only 'how to teach and develop curriculum' but also 'how to assess what we have taught!'"[32] Since 1970, the social work education literature has blossomed with offerings of relevance to our investigation. Unfortunately, few contributions have as yet dealt directly with examining and contrasting evaluation methodologies.[33] However, a good deal more preparatory discussion has taken place and more relevant studies have been reported upon.

One of the most useful of these contributions is Ronald Feldman's fertile article of 1972.[34] It is the most detailed adticle on evaluation in social work education yet to appear. After a review of Journal of Education for Social Work, Social Work, Social Service Review, and International Social Work, he concluded that very little critical self-evaluation had been undertaken in social work education and suggested at least eight factors which tend to inhibit the institutionalization of such evaluation.

(1) The evergrowing body of relevant knowledge militates against meaningful assessment of instructor competence.

(2) As specialization proliferates, both colleagues and students become decreasingly able to judge the competence and effectiveness of teaching specialists.

(3) Teaching competence is generally not so well rewarded as research productivity.

(4) The initiation of evaluative procedures is hampered by the very great work demands often placed upon faculty.

(5) While they do provide "indispensible safeguards" for educators, the tenure system and ethic of academic freedom tend to inhibit systematic observation and evaluation of teaching.

(6) Evaluation by other parties, especially by students, raises complex questions regarding expertise and professional self-regulation.

(7) Individual inadequacies and anxieties retard evaluative efforts.

(8) Unsatisfactory or ambiguous results from previous evaluative studies tend to discourage new endeavors.[35]

Although he does mention the need for more systematic techniques in observing instructors at work and does show considerable interest in rating instruments, Feldman's major preoccupation is with quasi-experimental approaches. He appears to make the assumption that "the harder the data, the better the evaluation." Of course, his concentration on objective measurement is a natural consequence of an overriding concern with assessing effectiveness. He suggests that the specification of criterion measures is a central methodological problem and takes note of the inherent complexity of such an undertaking when teaching effectiveness is viewed as a multi-dimensional phenomenon. He goes on to discuss various attributes and aspects of criterion measures,

and concludes by suggesting four primary directions for future research:

(1) "systematic examination and synthesis of previous evaluative research";

(2) "conceptualization of requisite theoretical and philosophical frames of reference";

(3) "formulation of basic research objectives";

(4) "development of appropriate methodological tools and techniques".[36]

Within its somewhat limited parameters, Feldman's article provides a very useful discussion about many important issues surrounding evaluations of instruction in social work. He pays quite considerable attention to: institutional barriers to the evaluation of social work teaching, the relationship between teaching evaluation and teaching proficiency, the delineation of criteria for evaluation, the selection of outcome measures, the differential efficacy of evaluations by colleagues and students, and future directions for evaluation research concerning teacher effectiveness. However, in restricting his discussion to validation studies, he deals with only a segment of the evaluative area. He included little or no consideration of effects other than those which are pre-specified as "criterion measures" and did not begin to explore the concept of "adequacy" or the levels of resource expenditure necessary to the conduct of evaluation research. Though we might conclude that the article was a commendable effort which has significantly enlightened quite a number of social work

educators, there would also appear to be a basis for the fear that it has inadvertently contributed to the development of an overly-narrow mind-set on the part of many who are interested in the evaluation of social work instruction.

It has already been noted that an increase in relevant critical comment began to appear in the professional literature early in the 1970's. Some of these commentators proposed conceptual formulations of curriculum planning and/or evaluation, while others simply expressed their reactive opinions. Charles Levy's "framework for planning and evaluating social work education" was a very sketchy "framework" indeed; but his article did explore some interesting, possibly useful, and certainly contentious points.[37] He began by noting that partialized approaches to modifications in social work education risk the achievement of one set of desirable objectives to the detriment of others; thus the need for more global consideration. He then proposed what might be labelled a "conservative" viewpoint toward curriculum change: ". . . it makes little sense to change a practice that has never been truly tried. It is only when a practice has received optimal application and attention that its outcomes become a relevant basis for change."[38] The operational word is, of course, "optimal". When has "optimal application and attention" taken place? This is the cutting point, and the exact location of the cut must depend upon who wields the knife. Levy fails to explore this point. Similarly, he fails to note that there may be a number of excellent reasons for a practice never to have been "truly

tried". If a particular practice has been available for a long time and yet is judged to never have been truly tried, then one might begin to suspect that there are plausible reasons. Perhaps the lack of "true" application indicates an innately perceived lack of potential which should be investigated before vast amounts of resources are committed to "true" applications. Be that as it may, Levy's major contribution is recognition of the fact that there are numerous, diverse, and often conflicting viewpoints which may be used to guide planning and evaluation in social work education. He repeatedly makes the point that, ". . . adequate planning and evaluation of social work education require a delicate balance among the interests of society, profession, agencies, practitioners, and clients, to all of which the school of social work owes some accountability."[39] He neglects to mention that the school might also bear some accountability to its students; but his general point should be a sobering thought for any curriculum developer or evaluator who undertakes to investigate even a small segment of an instructional program in social work.

Walter Walker does not propose any comprehensive conceptual framework for developing social work education. He does, however, note that social work education is embroiled in quite a number of basic conflicts, and suggests that this is a healthy state of affairs. One of the basic conflicts to which he alludes is a continuing controversy over "product" and "process". In his view, "There is too much emphasis placed upon process rather than product on the part

of the professional schools. Although it should be obvious that professional practice and professional education involve a consideration of ethics and process, I have always believed that our profession and the schools that educate for it have been more concerned with teaching techniques than with the question of whether or not these techniques actually work."[40] As has been noted here, the literature would appear to bear this out.

Walker goes on to explore conflicts around the granting of degrees, the objectives of social work education with regard to the screening and socializing of recruits, the utility of awarding grades, and the thought-provoking suggestion that it is not so much good teaching techniques as basically valid educational aims which are lacking in social work education. This last suggestion is supported by a somewhat startling quotation from Donald P. Hoyt.

> In a medical school it is certainly reasonable to expect that a graduate would be a better medical practitioner than a person who has not attended medical school. Are we as confident about the differences between a graduate of a school of social work and a practitioner who never metriculated to one of our institutions? It is probably true that a social work graduate would follow a set of guidelines and procedures in his work with a client. The available research does not make me as confident that this would automatically produce a better outcome from his client's point of view.41

In New Ways of Teaching Social Work Practice Catherine Papell reintroduced arguments for a unitary approach to the social work curriculum.[42] She argued that no part of the curriculum can stand on its own and that changes in one segment necessarily affect other segments. She further noted that a number of historical trends appeared to be

converging in the late 'sixties and early 'seventies to require a redefinition and reorganization of social work services and social work's educational complexes. In response, she proposed a set of basic social work practice skills which would serve to underpin renewed attempts at curriculum development, and she refined from them a relatively specific set of "learning objectives" which reflected a unitary conceptualization of social work practice. The resultant eight learning objectives are worthy of thorough consideration by any agency endeavoring to initiate or evaluate a comprehensive social work curriculum.

In the same volume, Lola Selby revealed the "Generalist Model" for teaching social work methods which was being developed at the University of California.[43] She emphasized, however, that each school of social work must find its own model of teaching, and that any school's particular model should be based upon the practice resources available in its region, as well as the particular demands of the profession in the areas served by the school. She also warned that the fieldwork component of any social work curriculum must be structured to reinforce the classroom teaching (or vice versa), otherwise fragmentation and disjointed effort are likely to result.

Finally, Jack Stumpf undertook to discuss a number of assumptions which underly the San Diego State College curriculum.[44] This included a presentation of perceived trends which tend to support the need for more holistically-based curricula. It concluded with an outline of fifteen

professional social work roles and an examination of some of the special educational needs implied by assumption of those roles. San Diego State, he said, was attempting to build an evaluative component into its program, and was encountering a major difficulty in that: "The changes in service delivery systems and the addition of new practice roles for social workers may lead to reformulation of practice before we can 'live with it' long enough to evaluate it."[45] Perhaps this is a rather gross manifestation of the concerns which led to Levy's "conservative" statement mentioned earlier.

The reader will have noted that the foregoing few pages have dealt, primarily, with contributions to the social work literature which emphasize curriculum development issues but are, of necessity, intimately involved with the subject of evaluation. Other discussions and examinations of the same genre are, of course, well worth reading, but form too extensive and peripheral a collection for review here.[46]

In April 1973 a symposium on evaluation and accountability in social work education was held at the Graduate School of Social Work, University of Utah. Papers were commissioned in three areas: (1) "discrepant role expectations of social work education", (2) "present trends in evaluation and accountability", (3) "setting objectives as a primary step in evaluation and accountability".[47] Unfortunately for our purposes, very little attention was paid to the actual activities involved in evaluation and a very high proportion of

the time appears to have been spent on issues related to accountability. As is so often the case at such gatherings, the more difficult topics tend to be overlooked. Be that as it may, some material relevant to our enquiry did accrue from the symposium and, thus, a review of the papers is in order.

Gordon Hearn presented a paper in which he discussed the concept of "accountability" and some of the implications of accountability demands on social work education.[48] The tenor of his discussion was general trepidation and his intention appeared to be primarily in warning the profession of associated dangers. He expressed fear of an overemphasis on accountability; arising from characteristics of the concept itself.

> Because I believe that we achieve knowledge through feeling as well as thinking, through the exercise of the emotions as well as the intellect, I also believe that the preservation of spontaneity is terribly important. Perhaps this is what bothers me most about accountability. It suggests a terribly rational, intentional and deliberate process and I know, and you know, that a great deal of what we do is intuitive and much of what we know is the result of an accumulation of a host of unexpected events and experiences.49

He then suggested a number of potential hazards associated with demands for accountability: in the extreme it easily becomes repression; it can be a "red herring" and a distraction; it can be a disguised form of harassment and scapegoating. These considerations lead him to two warnings: (1) "As we strive for accountability I believe we should be acutely aware of who is demanding it and why. . . . We should never forget that we seem to tend, as a society, to

value tranquility more highly than we do freedom"; (2) ". . . the effective regulation of our work is best done by ourselves but we must take extraordinary efforts . . . to measure our effectiveness. If we don't others will, and they will do it poorly."[50]

In his final section, Hearn explores more completely the problem of "objective" measurability. He notes that there are two reasons why room must be left in evaluation efforts for the unmeasurable. First, he says, the intuitive side of teaching practice is valuable. "The second reason is that we may confuse for reality the things we see because of a particular perceptual set derived from the way we have defined our objectives. There is the danger that we may regard as desirable only those behaviors we have predefined as desirable and downgrade the unexpected and unconventional."[51]

In response to these comments, Richard Lindsay admits that there are many aspects of social work education which cannot be measured, but he suggests that we emphasize the things that we can measure, and he repeats the warning that if the profession does not pay much more attention to accountability measurements, outsiders will begin to make them instead.[52]

In a second paper, the executive director of the Council on Social Work Education points out the decreasing amounts of funding available to social work education in America and suggests that, increasingly, those funds which are available for innovation will be allocated to proposals

that have sound evaluation mechanisms built in.[53] Here the operative concept is "sound". Though he notes that specificity of objectives is very difficult to accomplish in a profession which is rapidly becoming more diverse, he makes it clear that his concept of "sound" evaluation requires it. He leads one to the somewhat pessimistic conclusion that evaluative investigations which do not produce observations acceptable to a "scientifically oriented" society, though they may be valid and useful studies, will be less influential; possibly a sobering thought for any investigators who refuse to adopt scientific positivism. His concluding statement, however, should provide a ray of hope for anyone committed to the practice of evaluation.

> I would hope that every school would designate one or two faculty members to be the recipient and creator of evaluation methods, to know the literature on educational measurement, to devise how that works in social work education, and that the Council [CSWE] become a repository and have the opportunity to pull together and to make available to the field of social work education what the evaluatory experience of various schools are--that is a role that I see for a national organization in social work education planning.54

Rex Skidmore's paper began with a brief restatement of the traditional "classico-experimental" paradigm for evaluation research. He expounded at some length on the need for behaviorally defined objectives, and suggested that their development was a major step toward "genuine accountability".[55] Skidmore then went on to propose five types of evaluation which might occur within schools of social work; (1) instructor self-evaluation, (2) student

evaluation, (3) faculty evaluation, (4) outside evaluation, and (5) administrative evaluation.

In commenting on Skidmore's contribution, Horace Lundberg castigated past evaluative efforts for choosing the most easily evaluated objectives.[56] He suggested the alternative of attempting to evaluate those objectives which we value highest. "Evaluation", he said, "to some extent relates to expectation. Someone has said that the minimum expectation becomes the maximum accomplishment. . . . The highest priority objectives then should be the center of our evaluation and we should be clear about our expectations." He advocated that instructors pre-specify their objectives with the learner but he did not attempt to explore possible modalities for the negotiation of objectives. Almost naively, one might think, Lundberg then notes that if we have our objectives "spelled out", we can measure whether or not we are accomplishing those objectives. But he is not entirely naive: witness his closing statements:

> There is a danger, too, in spite of the positive remarks made about behavioral objectives. The danger is that we may overemphasize them. We may let the evaluating structure be too controlling of what we do, then evaluation becomes the dictator of program . . . when we start evaluating and looking at our measurable objectives, we need to be aware that focusing only on that which is quantifiable raises a tendency to operationalize that which is easiest to operationalize.57

Thus far our review has included only one enquiry which relates the results of empirical research as such. Most of the articles are concerned with discussing the various issues related to evaluations of instructional programs in social work. Relatively few published reports

of evaluation studies in social work education have appeared in the literature, but their incidence does appear to be on the increase. Furthermore, a number of other published reports deal with investigations which might not strictly fit within the category of "evaluation" but which, nonetheless, are of interest to those contemplating the evaluation of an instructional sequence. My review of both sorts of research is by no means exhaustive, reports have been chosen more for their illustrative qualities than for any other reason, but it is hoped that examination of a few studies will provide an indication of the current state of evaluative research within the area.

It would seem appropriate to begin this section of the review with a simple quasi-experimental study. A good example is the report on evaluation of the C.A.R.E. Project; an educational program involving medical, social work, and nursing students.[58] The major educational objective of the project was to "positively affect student attitudes toward the aging" through interdisciplinary team participation. Very simply, the twenty-four students in the program were compared with twenty-four "control" students who were matched for age, level of training, and profession. Osgood's Semantic Differential was used to measure attitudes toward old people, nursing homes, death, aging, and chronic disease; while Kogan's Attitude Toward Old People Scale provided measurements of stereotyped thinking about old people.[59] Both scales were administered as students entered the project and again after six months. Data then underwent

analysis of covariance to determine differences in the degrees of change between "treatment" and "control" groups. Briefly, the results showed that students involved in the interdisciplinary team project developed significantly more "positive" attitudes toward aging, old people, and death than did the "control" subjects. There was also some indication that the experimental group members developed less "stereotyped and negative" attitudes on certain items contained in the second scale.

This relatively straightforward study displayed a number of positive points: it made a significant attempt to use standardized measurements of attitude; it assessed students prior to involvement in the program in order to establish a baseline against which change could be inferred; it used more than one assessment instrument; and it included a matched control group. Unfortunately, it provided scanty information about the instructional program to which students were exposed and it did not specify what the control group subjects were doing during the six month interval. We are not informed of the extent to which control group subjects were exposed to old people, nursing homes, death, and chronic disease. In fact, from the limited information given, it is impossible to reach any substantial conclusions about the genesis of the apparent differences found. Inclusion of the control sample appears to have done little. Certainly, the study provides no indications with regard to the adequacy or effectiveness of alternative instructional modalities, and thus, no real basis for a

comparison. It might also be noted that the study does not offer any breakdown of data to differentiate between medical, nursing, and social work students. Thus, we are left with no idea of whether any changes that might have occurred were homogeneously spread throughout the group or differentially apparent in some way.

A similar research design was used by Kolevzon in a study which assessed the effectiveness of a "teaching center" model for intergrating the research component with other learning units in an M.S.W. program.[60] Kolevzon's report is much more thorough than the previous one and a brief summary is unlikely to do it justice. Unfortunately, our summary must be brief. The examination begins with a discussion of the concept "integration" in social work education which explicates two somewhat dichotomized "perspectives" on the concept: "organic" and "mechanical". Descriptions of the "experimental" and "contrast" groups used, and a brief outline of the aims of the "teaching center", follow. He then poses the central question which the study is designed to answer.

> How did the students' attitudes toward research and their actual learning about research methodologies within an experimental (integrated) research seminar compare with those students in a concurrent research seminar (contrast group) taught by the same instructor, but one in which there was no explicit attempt to structurally integrate the students' research course with other learning units in their second year program?61

Kolevzon notes that both pre-tests and post-tests were used on assessments of "attitudes toward the consumption of research" and "ability to describe and critically assess a

research case study". He discusses both the double attitude scale and the cognitive test used to assess change, and reports on the study's findings: contrast group students rated consistently higher on both scales of the attitude assessement; as gauged by the cognitive test, contrast group students learned more. The excellent discussion of alternative hypotheses, generalizability, instrumentation, and implications which concludes the report makes it stand out as something of an exemplar.

Another somewhat similar study was undertaken by Sharwell.[62] This investigation sought to answer two questions: ". . . does the orientation toward public dependency of students entering a graduate school of social work differ significantly from an undergraduate population? and ". . . does the orientation of students toward public dependency change over the course of the two academic years of graduate social work education?". It used a simple pre-post-test design and a Likert Scale to measure attitudes toward public dependency (n = 20). Some validation work was done on the instrument, and comparison was made with faculty attitudes measured on the same scale. The findings suggest that students entering a graduate program in social work display an orientation to public dependency which is more compatible with the public dependency orientation of professional social workers than could be predicted from the scores of an undergraduate population. Sharwell hypothesizes that this is a result of both student self-selection and school selection and admission procedures. The second

finding suggests that students' attitudes did change significantly over the course of their studies. It supports the conclusion "that graduate social work education does influence student orientation toward public dependency in the desired direction."

Since his findings tend to contradict those of previous similar studies, Sharwell enters into a discussion of possible reasons for the discrepancy.[63] Though far from exhaustive, it is somewhat provocative. His tentative conclusion, in which he attempts to use "cognitive dissonance theory" to support his findings, is interesting but not well supported.

Sales and Navarre evaluated group supervision in social work education, to determine: whether or not it saves supervisors' time, and whether student progress suffers.[64] They used matched samples of first and second year MSW students; the "experimental" group receiving group supervision, the "contrast" group receiving individual supervision. Questionnaires were administered to both students and supervisors, and comparisons were made with regard to reported use of time and differential views on learning achieved. It was concluded that group supervision did save supervisors' time and that there were some noticable differences in practice skill development between the two groups. However, no conclusion about the overall superiority of one supervision method was advanced because specific criteria for judging them had not been set out. Studies of this sort might be looked upon as much more exploratory than

experimental. Attempts were made to match the two groups, but the data gathered was highly subjective, no pre-test component was included, and samples were small.

A larger and somewhat more adequate quasi-experimental study was undertaken by Goldstein and Horder who studied 220 students taking introductory research courses at five American schools of social work.[65] The researchers used Measurement of Attitudes and Research Knowledge (MARK) to assess student attitudes and research knowledge before being exposed to the introductory research course, after that course, and just prior to graduation.[66] Data was analyzed with regard to information provided by the twenty-one course instructors involved. Instructor characteristics, learner characteristics, and characteristics of the instructional program were arranged into "patterns", and conclusions were drawn from comparisons of results. For example:

> . . . students with a good background in research learn best when a variety of content is introduced; students with poor backgrounds gain most with limited content. All students appear to learn best when taught by didactic methods and evaluated by objective evaluations as compared to other methods. Differences in teacher characteristics such as age, field of study, and experience had little impact on student learning.67

The apparent object of this research was to develop a "suggested model teaching plan" for introductory research courses which would take account of various instructor-related, content-related, student-related, assessment-related, and teaching-related variables. Although definitely a worthwhile undertaking, the study demonstrates the need for very large samples in research of this sort. For, although

229 subjects were used, the various sub-samples occasioned by the division into "patterns" were relatively small, thus indicating a need for much caution in interpretation.

A study by Thomlison and Seidl reflects a somewhat different approach. They utilized three distinctly different information-gathering media to evaluate an innovation in educational programming, ". . . wherein the students would be free to identify their own educational needs and objectives, and to pursue their learning both individually and in the context of small group interaction within a 'problem framework.'"[68] The "problem framework" was provided by a six page case example which incorporated twelve current issues in social work. It was intended to be used primarily for definition of content boundaries, as a focusing device, and as a point of departure for group discussions.

The first evaluation technique involved "construction and administration of an equivalent forms pre - post multiple choice test of social work and social welfare knowledge." This procedure reflects the now familiar quasi-experimental approach to evaluation research; but, like many others, it included no "control" or "contrast" groups against which to compre findings. Results indicated that the average student improved his score on the knowledge test by slightly more than 10 percent.

The second evaluation technique is somewhat more innovative. At the beginning of the instructional term, each student and faculty member was asked to state his own

objectives for the course of studies. In addition, each seminar group was asked to formulate a statement of group objectives. At the end of the instructional term, these sets of statements were returned to the participants and each was rated on a seven-point scale, the poles of which were labelled "objective not at all achieved" and "objective completely achieved." Provision was also made for changes in objectives during the course of the program, as well as explanations of those changes. The findings were summarized as follows:

> The ratings of all the groups and individuals involved averaged above the three-point neutral level on all counts. One third of the students elected to state alternative goals and their attainment of them. Their ratings of attainment on alternative goals was substantially higher than their ratings of the initial goal statements. . . . This was not the case, however, when faculty chose to change their objectives. . . . Only two of the nine instructors stated changes in objectives, however.69

The third evaluation technique involved formal interviews with all nine instructors, three weeks after the close of term; i.e., after program termination. The report does not provide much information about the dimensions of the interview schedule used, but it does suggest that instructors were given an opportunity to explore and explicate their reactions to the instructional program in a relatively unstructured manner. On the whole, instructors appeared to favour the innovatory program over a return to the standard curriculum which had been used in previous years. This was, however, a majority agreement; not a unanimous choice.

This study is of particular interest because it attempts to combine three relatively dissimilar evaluation techniques into one investigation, and because the results indicate substantial agreement between the findings of all three. Unfortunately, none of the procedures are described with attention to details of application and analysis of results. Furthermore, no attempt appears to have been made either to examine parallels and divergencies between the results of the various techniques, or to explore significant differences between the nine seminar groups studied. In other words, a much broader analysis of the data could have been undertaken.

A study of the University of Wisconsin's "3-2 Program" by Alfred Kadushin and George Kelling is by far the most compellingly admirable investigation of its sort discovered in the social work literature.[70] Very briefly, the objective of this enquiry was to compare educational outcomes of a group of students who had received three years of undergraduate and two years of graduate education with the educational outcomes of a similar group of students who had received the traditional four years of undergraduate and two years of graduate education. The investigators were interested in determining "the extent to which the 3-2 group was the same as, or different from, the 4-2 group in terms of social work knowledge, attitudes, and skills at the outcome of training."[71] The design, therefore, called only for testing at the end of the instructional period. What distinguishes this study is the scope of the

evaluation measures employed. A sixty-item information test, the Tulane Assessment Scale of Casework Knowledge, and the Benjamin Rose Institute Interrupted Case Test were used to assess cognitive functioning. Myers' Social Work Values Test and a self-awareness measure assessed attitudinal states. An audio-taped work sample and autocritical analysis, a client reaction form, peer ratings, and standardized field instructor ratings formed a composite assessment of skill in performance on the job. In addition, four "global" measures were used: (1) grades achieved, (2) review of differential dropout rates, (3) a monthly log completed by students, and (4) a follow-up form soliciting performance assessments from the employing agency one year after graduation. In all, 98 specific outcome measures were spread over the various procedures used. For the record, findings indicate quite conclusively that graduates of the 3-2 experimental program were indistinguishable from those of the 4-2 program, on levels of social work performance. Undoubtedly, these results contain significant implications for traditional social work education, but for our purposes the investigators' thorough discussions of outcome measures employed are the major attraction. Almost one third of the two hundred page final report is concerned with describing the measures used and exploring numerous associated issues, problems, drawbacks, and implications.

Finally, interesting recent work has been noted by Tully who introduces Personal Construct Theory and examines the area of psychological change related to professional

social work training.[72] He reviews a number of recent studies which have examined aspects of psychological changes occasioned by social work training, using a repertory grid technique.[73] He does draw a few implications for social work instruction from the findings, and suggests that, with refinement, such an approach may provide us with a very useful evaluative technique.

We have noted that the conventional approach to research in social work education has assumed a positivist orientation, characterized by classico-experimental studies. This stance is criticized by Carrier and Kendall who suggest that it tends ". . . to ignore the processes involved in the creation and sustaining of social reality . . . and to disregard the existence of an infinite variety of social realities . . . and of infinite interpretations of any one social phenomenon."[74] They suggest that such an approach assumes that social phenomena are for all analytical purposes qualitatively the same as "natural" phenomena, and thus are subject to similar techniques of analysis. However, they note that what one person conceives as rational, logical and practical may not be so conceived by another; thus, that the positivist orientation ignores the centrality of value judgements in analyzing social phenomena. They go on to suggest an alternative, labelled the "phenomenological perspective", which examines questions like "how did such and such come about" and "what has such and such accomplished" by studying the processes through which any body of knowledge comes to be "socially established and sustained as reality."

Phenemonology, they say, places primary emphasis on how members accomplish social interaction, and is based upon an acceptance of the presence of multiple realities. Rather than acceding to the demands of natural science hypothesis testing which forces data into preconceived "objective" reality, they advocate that theories should be ". . . derived from the data gathered and in particular make use of the categories that participants themselves use to order their experience."[75]

Joseph Vigilante relates the same criticisms and injunctions even more closely to research on instruction in social work.[76] He contends that:

> As social work has grown during the ascendency of science in the Western world and during the objective-analysis emphasis in the social sciences, it has really postured, at times gyrated and stumbled across the intellectual and scholarly stage to create its own theatre of the absurd. Like the social science upon which it depends for much of its knowledge base, social work has attempted to ape the natural sciences.77

In contrast, he advocates a more "engaged style" of scholarship and suggests that social work educators will find it more highly compatible with their own educational theories and practices.[78] "Indeed," he says, "it is the engaged style of social work education that traditionally has been contrary to twentieth century 'scientifically oriented' education, and which I think we have been neglecting in favor of 'scientific' approaches. Ironically this style may be social work's most important contribution to higher education."[79] Finally, he warns that the adoption of this allegedly more compatible style of scholarship may lead to

accusations of anti-intellectualism and he refutes such indictments:

> It is after all fallacious to confuse scientific endeavors with intellectual endeavors, or to assume that the scientific is necessarily scholarly or intellectual. Logical positivism is appropriate and highly useful when it fits--when it investigates distances, shapes, depths, temperatures, speeds, mass, density, and other physical aspects of the universe. It has not been proved to be efficient for understanding those processes, actions, and behaviors associated with human relations.80

As might be expected, examples of the exclusive use of this more "engaged" style of scholarship in the sphere of social work education are rather rare. Goffman and Whyte have contributed major studies in other substantive areas which utilize such an approach and will be familiar to most social workers.[81] More directly related to educational concerns, similar studies have been carried out on education for the medical profession and forms of student assessment.[82] Of direct relevance to instruction in social work is the investigation by Brennan and Arkava, which initially appears to reject the traditional approach.[83] In this study the investigators attempted to construct a student-based evaluation of the undergraduate social work program at the University of Montana by interviewing all senior year social work majors registered for the 1973 winter quarter. They argued that regardless of the outcomes which are intended for an instructional program, the "raw material" (i.e., students) largely determines the products produced.[84] Brennan and Arkava's reliance upon the extended interview as a research tool suggests that their investigation rejected a quest for quantified, empirical data and accepted a more

phenomenological approach. A closer reading, however, indicates that the investigators did not really undertake a study based in a "grounded theory" approach, rather they conducted interviews in an entirely prescribed manner. They asked only a series of pre-specified questions which corresponded to certain initial assumptions made by the research team.

> Did they prefer a conceptual or a pragmatic approach to learning practice skills? How did they differentially regard the six program objectives set forth in the CSWE guidelines? What specific required courses within and without the department did they find useful? What were their thoughts on specialization at the undergraduate level? To what extent was learning in the field preferred to classroom learning? How useful did they regard simulation techniques in acquiring practice knowledge and skills? And, finally, did they find the professional journals in the field to be at all informative or useful.?85

It would appear, then, that the study was not based upon any conviction that the findings might be structured largely by the perceptions reported by their subjects; rather, it further promoted an approach in which findings are structured by the pre-conceived notions of the researchers themselves. It is not argued that a pre-specified interview pro forma necessarily pre-determines details of the results obtained, only that exclusive use of, and rigid adherance to, such an instrument strongly predisposes the results toward compliance with given perspectives. In relation to the previous quotation, witness the study's "Summary of Findings."

> The average respondent may be described as a young, white, single female who had decided on her major early in college and had been disappointed in didactic coursework because it was not practical enough. She wanted to understand and help people and was looking for practical knowledge from college

> that could be applied to real situations. She found actual practice with supervision the most helpful way to learn. She also liked simulation techniques and favored early specialization.86

This investigation does not appear to have uncovered one significant finding beyond what it intentionally set out to find. While it reaches a conclusion like: "This is obviously a poorly structured model to enhance learning",[87] the study makes no effort to investigate the perceptions and attitudes of other significant individuals and collectives (e.g., instructors, administrators, clients, practicing professionals, specialists in education). Further, it undertakes absolutely no investigation of "learning"--only student reactions to their particular experiences in that particular school, as elicited by an entirely pre-conceived data-gathering instrument. Thus, such a sweeping conclusion is not at all well substantiated. At best, such a limited study could validly produce only very limited, tentative conclusions about students' reported reactions to a specific and rather limited set of researcher-conceived issues.

A study which much more validly represents the principles of an "engaged style" is provided by Gillian Michael.[88] This is the only located investigation of aspects of social work education which substantially adopts such an approach. Michael notes that her study is "essentially descriptive and interpretative." She utilized participant observation, formal and informal consultation with participants, interviews, questionnaires, and documentary sources to "triangulate" on the experience of fieldwork instruction as perceived by fieldwork teachers, course tutors, and students. Rather

than attempting a study based on "objective measurements" of student learning, she developed a detailed portrait of the fieldwork experience which included consideration of myriad important aspects of process and structure that emerged during the course of her investigation. That is, she attempted to tailor her investigative techniques to the people and milieu under study, rather than vice versa. In this way, the viewpoints of students, tutors, and fieldwork instructors assumed a prominent role in creating and ordering the set of images produced. Of her data collection methods, she says:

> Data was collected from a wide range of sources, meetings, interviews, formal and informal interactions, and in a wide range of places from classrooms to coffee rooms to pubs to that never failing source of material--"the Ladies". Use was made of semi-structured interviews and of questionnaires. The latter were to a large extent unsuccessful, the product of my inexperience I fear, for silly questions get silly answers.89

A colourful and detailed picture of the fieldwork experience, viewed from a variety of perspectives, was the result. However, Michael did not suspend her investigation at that point. She went on to deal at length with explanations, explorations, and criticisms of the study's adopted investigative methods, more obvious weaknesses, and inherent problems--both in theoretical and practical terms. The overall result is, therefore, not only an edifying "illumination" of the chosen subject matter, but also a highly instructive document for any investigator contemplating the use of a phenomenological approach in exploring aspects of social work education. Michael's sensitive yet forthright

observations with regard to the numerous problems, issues, and constraints which arose should be of guidance and considerable consolation to those undertaking similar investigations.

The final study chosen for this review takes an approach to evaluation which fits comfortably within neither the "illuminative" nor the "experimental" category. Edward Brawley utilizes a "systems framework" for analyzing an instructional program for non-professional mental health workers.[90] He describes some antecedent conditions which led to institution of the instructional program, outlines the program's objectives and design, and proposes an analytic scheme which ". . . utilizes a general systems approach to program structure and performance, incorporating data obtained from statistical records, from program personnel, enrollees, and interested others, and from observation of the program functioning. Quasi-experimental techniques are used to measure the effectiveness of program in accomplishing specific subgoals."[91] A "table of organization" and a "flow chart" serve as a point of departure from which the analyst explores program structure and performance. The outline below portrays the various divisions used to accomplish that exploration.

I. Structure

- A. Human Resources
- B. Nonhuman Resources
- C. Subsystems
- D. Internal Relations
- E. External Relations
- F. Values
- G. Guidance and Control

II. Performance

A. Satisfying Interests
B. Output
C. Investing in the System
D. Using Input Effectively
E. Acquiring Resources
F. Observing Codes
G. Behaving Rationally

The paper also includes some consideration of issues and problems that accompany the use of existing records, observational data, interview data, and pre-test/post-test procedures. Taken as a whole, it pursues an approach to evaluative investigation which utilizes elements from both the "quasi-experimental" and the "interpretative" paradigms, but wholly embraces neither. Brawley concludes with a short discussion of the major virtues and dangers involved in using such an approach, but does not evolve these considerations to any real depth.

This brief review of the literature has indicated that until very recently only a small number of investigations of instructional practices in the field of social work appeared in the professional literature. In summary, almost all of the earlier contributions were theoretical explorations which may be regarded as "scene-setting", but neither descriptive nor analytical studies. The bulk of the empirical research which has been done in this area adopted quasi-experimental designs, concentrating attention on the assessment of program effectiveness. The use of alternative designs to explore different perspectives and utilize non-quantitative information appears to have had a severely retarded development.

CHAPTER III

THE PARADIGMS

The relatively extensive and recent literature on evaluating instructional programs suggests that accelerating demands for evaluation and accountability in educational spheres is being keenly felt and that evaluation is fast becoming a special area of competence within the profession of education. We have noted previously that within this literature there has developed a rather loosely knit set of prescriptions and activities which are used to determine the merit of instructional programs. In order to examine it further, the field might be structured in a number of ways, for example: by chronological exposition of methodological developments, through classification of methodologies according to their disciplinary biases, by identification of "schools". I intend to begin by outlining two highly generalized evaluation paradigms, or ways of thinking, which guide evaluation efforts and differ considerably from each other. In this Chapter it is my intention to examine these paradigms in some depth and contrast them, so as to lead us closer to a useful comparison, at least in theoretical terms.

N. L. Gage has noted that paradigms are not theories, rather they are "models, patterns, or ways of thinking or patterns for research that, when carried out, can lead to the development of theory."[1] They are highly generalized constructs and their usefulness derives

precisely from that level of generality, in that they may be applied in a very large number of specific instances of a whole class of events or processes. Obviously it is important that we pay considerable attention to the paradigms which guide evaluation research for, as Gage says,

> When one has chosen a paradigm for his research, he has made crucial decisions concerning the kinds of variables and relationships between variables that he will investigate. Paradigms for research imply a kind of commitment, however preliminary or tentative, to a research program. . . . Choice of a paradigm, whether deliberate or unthinking, determines much about the research that will be done. The style, design, and approach of a research undertaking, indeed, the likelihood that it will bear fruit, are conditioned in large part by the paradigm with which the investigator begins.2

I have suggested that there are at least two major paradigms which are being used in the evaluation of instructional programs. Others would argue that there are more, depending on the sorts of dimensions examined. If evaluation methodologies are analyzed with regard to their "focus of attention" (i.e., what are their primary foci?), one notes that one set of methods appears to be based upon a major interest in the outcomes or products of the instructional program, while another focuses much more heavily on the processes which occur within the instructional program. Though both paradigms may be used in examining the same instructional programs, they tend to focus attention differentially on either products or processes. Furthermore, some of the assumptions upon which each is based tend to make one exclusive of the other. If these product-process orientations can be viewed as not entirely mutually exclusive,

one might begin to suspect that they exist more on a continuum between two extremes, rather than simply as discrete conceptual entities. Because of their very high level of generalization, the paradigms overlap with each other considerably in application, and the areas of overlap form the central segments of the continuum.

It is suggested that the paradigm characterized as the "classico-experimental" approach more closely approaches the outcome-oriented pole of this continuum, while the paradigm labelled the "socio-anthropological" approach more closely approaches the process-oriented pole.[3] Although this division tends to break down to some extent when the paradigms are actually applied in specific evaluation studies, examination of their "pure forms" should provide us with sufficient understanding to begin to assess their relative strengths, uses, and weaknesses in theoretical terms. At the same time, much of this discussion must be viewed as a presentation of claims, counter-claims, and hypotheses about the two paradigms. The presentation of evidence to support or reject these claims and hypotheses can only be undertaken in an extremely limited fashion. Highly generalized constructs such as these must be tested and observed repeatedly in numerous forms of application before any really firm conclusions can be drawn. This thesis will be considered a success if it manages to begin that ball rolling within the domain of social work education.

Classico-experimental Paradigm

Approximately four decades ago Ralph Tyler and associates began to publish a number of works which advocated and outlined the development of a technology of curriculum which concentrated attention on a matching between observed behaviors and stated objectives.[4] His approach was germinal to the vast majority of curriculum development and evaluation work which has been undertaken since that time. Basically, Tyler suggested that if one wished to assess the effectiveness of an instructional sequence one needed to begin by operationally defining instructional objectives; only then could one concentrate on taking the appropriate measurements. He stressed, therefore, how important it was that the criteria against which a course's effectiveness was to be assessed should be specified in measurable terms. In this regard "objectives" are not coterminous with "goals" or "aims"; "objectives" constitute operationally defined answers to the question, "What will the student be able to do as a result of the teaching that he was unable to do before?"[5] The thrust of this approach to evaluation is succinctly summarized by Jones and Borgatta, "If the purpose of evaluation research is the determination of the accomplishment of program goals, these objectives must be clearly formulated from the beginning or their measurement is impossible."[6] Once objectives are operationally defined, it becomes possible to use an hypothetico-deductive approach to data generation and theory-building. In pure form, the classico-experimental approach

to evaluation advocates the following steps:

(1) broad instructional aims are itemized and qualified until they represent specific, operationally defined learning objectives which state what the learner will be able to do, or do better, as a result of experiencing the course of instruction;

(2) whenever possible, a "treatment" and a "control" group are designated; subjects are assigned to each, preferably by random allocation;

(3) instruments are developed to measure individual performance against the criteria outlined in the objectives;

(4) the instruments are applied to measure each individual's performance (pre-test);

(5) "treatment" group members are exposed to the instructional program while "control" group members are not;

(6) instruments are applied again to measure the performance of all individuals (post-test);

(7) data from both instrument applications are analyzed and compared between pre-test and post-test, and between treatment and control group members;

(8) conclusions are drawn from the analyzed data.

Obviously, then, the paradigm emphasizes measurement of program outputs in order to ascertain the extent to which learning objectives have been met. A number of points should be made about each of the outlined steps.

Specification of Objectives

Operationally defined objectives are not dictionary definitions. That is, they are not defined in isolation, but by stating the observable conditions under which they operate. As Brodbeck puts it, "In general, definition in science stops when all descriptive terms in the definition refer either to some physical objects, or to some directly observable properties and relationships of and among them."[7] Within the sphere that concerns us here the "objects" are usually people while the "directly observable properties" are usually the behaviors of people, i.e., all human activities that can be assessed or inferred. Operationally defined learning objectives, then, are statements of intent about what the learner will be able to do, or do better, as a result of instruction and are defined so that learners' attainment of these objectives may be empirically assessed. Well defined objectives tell us precisely what to look for in the student's behavior.

The educational research literature is replete with discussions of objective development and definition.[8] The gist of these discussions revolves around the assertion that instructional programs are never aimless. We noted in Chapter 1 that each "actor" in the process has relevant purposes, goals, or aims which he attempts to maximize. Such aims provide both the standards against which objectives are justified and the material out of which objectives are created. The definitional point at which vague "aims" become specific "objectives" is not, however, clear. "There

is no profit to be had from trying to stipulate at what point of precision a statement becomes an objective rather than an aim. (When does fog become mist become clear visibility.). The essential thing to recognize is that the nearer we get to an objective, the more likely it is that 'interested parties' or 'experts' will agree as to whether or not a given student's behavior indicates attainment."[9] The point is that we must move toward this higher level of specificity in order to attempt objective assessments of program effectiveness.

If each participant in the instructional process has aims, and consequent possibly-identifiable objectives, there are bound to be a variety of objectives for any particular instructional program. This point was mentioned earlier in one context. It should be noted also that such variation can be observed in the kinds of behavior implied by the objectives. For example, an instructional sequence may be aimed at acquisition of a certain body of knowledge, integration of particular affective responses, or the development of some practical skill. The two major taxonomies of educational objectives attempt to deal with the "cognitive" and "affective" areas respectively.[10]

> Cognitive: objectives which emphasize remembering or reproducing something which has presumably been learned, as well as objectives which involve the solving of some intellective task for which the individual has to determine the essential problem and then reorder given material or combine it with ideas, methods or procedures previously learned. . . . Affective: objectives which emphasize a feeling tone, an emotion, or a degree of acceptance or rejection. Affective objectives vary from simple attention to selected phenomena to complex but internally consistent qualities of character and conscience.11

These detailed taxonomies also attempt to distinguish between objectives with regard to their level of complexity; i.e., some objectives are more difficult to attain, often because the learner must first acquire a number of lower-level ("enabling") abilities. The taxonomic schemes may have some value as one conceptualizes and redefines learning objectives, but they also have their deficiencies: (1) there is no encompassing classification of the "skills" domain; (2) there are large areas of overlap between the levels of complexity; (3) the examples quoted are often not clear; (4) and their form tends to support an implicit assumption that cognitive and affective objectives can be attained independently of one another while it would appear that most "real-life" learning objectives are an amalgam. As Rowntree puts it, "We are not interested merely in what the learner can do (immediately the course is over) but, more importantly, with what he will do (when the course is long past and he is free from the threat of assessment). It is the affective element that turns 'can do' behavior into 'will do' behavior."[12] The distinction between "cognitive" and "affective", though possibly valuable in analytic tasks, suggests an ordered vision of reality which may seriously distort the essentially unitary nature of educational goals and outcomes.

In summarizing the benefits to be had from operationally defining objectives it is suggested that they not only (1) help to ensure that methods of evaluation are objective and appropriate, but also that they (2) help us

to select and structure instructional content, (3) assist us in deciding on appropriate teaching methods, and (4) enable us to communicate more clearly about our intentions.[13] The first three points are self-explanatory; the fourth requires further elucidation. It asserts that specific definitions of learning objectives enable the instructor to communicate more clearly with his students; with colleagues, administrators, other interested parties; and with himself. One value of this point is highlighted in a study by Mager and Clark who found that when students know the objectives of a course, and are given appropriate references and resources, they can teach themselves in half the time taken by more "normal" classroom methods.[14] Further, the attempt at objective specification can serve as an excellent vehicle for initial negotiations between students and instructors, while discussion with colleagues, administrators and others creates an opportunity for dispute, argumentation, and cooperation. "Broadly phrased curriculum and course aims usually sound far too worthy to quarrel with. Yet when they are translated into objectives the illusion of consensus is often shattered. Only then can teachers enter into a productive debate about what exactly is to be taught and what are the implications for the system as a whole of the objectives that different teachers want achieved."[15] Finally, it is suggested that the self-awareness produced by clarification of teaching intentions leads the instructor to more purposeful teaching and the student to more purposeful learning.

Treatment and Control Groups

Adoption of the classico-experimental paradigm requires that, whenever possible, a "control" group be used in order to generate evidence that it actually was the instructional sequence, and not some other variable(s), which "caused" the manifest outcomes of instruction. As with other experimental research, two groups are composed (most often randomly drawn) from a pool of possible subjects, the "treatment" group receives instruction while the "control" group does not, and post-test performance measures or indices of change for the two groups are compared. This may provide an effective means for substantially ruling out the possibility that influences other than those of the program caused any improvements that were observed. More basically, if the treatment group members' performance changed in the anticipated direction significantly more than did control group members' performance, the procedure may yield strong indications that the instruction played a large part in producing the change. If changes between the two groups are very similar, we have an indication that the instruction did not produce significant appropriate changes. Discussions of control-treatment group theory and technique can be found in any textbook or experimental research and need not preoccupy us here. What should be noted, however, is that control groups are very often not practically available in evaluative research on instructional programs. Very often the only subjects available for study are those who will be exposed to the instructional effort.

Where control group members simply are not available, resort is sometimes made to a "comparative design", but in a great many instances, practical or ethical considerations prevent the use of "control" or "contrast" groups even if they are theoretically available.[16] Of necessity, this point will arise again later.

Development of Instruments

In Chapter I we discussed a number of media, or instruments, which could be used in evaluating a program of instruction. The classico-experimental paradigm appears to focus primary attention on the use of relevant tests--performance measures which are based on the program's instructional objectives. Of course, combinations of various sorts of instruments might be used in any particular evaluation study, but the classico-experimental paradigm's emphasis upon assessment of behaviorally defined objective achievement heavily supports the use of tests which are as objective as possible. Bloom suggests that there are three major reasons for using achievement tests: (1) to assist in determining whether the teaching method, instructional procedure, or teacher does produce changes in learners; (2) to identify samples of students being studied and assess whether there are differential effects of teaching on the various sub-groups; and (3) to determine how changes in one domain of objectives are related to changes in other areas.[17] We have already noted that a great deal has been written on test construction and social science measurement in general.[18] We should, however, take note of an important

distinction between two sorts of tests. In social science generally, tests are designed to measure differences between individuals. The test constructor attempts to focus the facility value of items on a norm, often the fifty percent level of attainment. For this reason, these are called "norm-referenced" tests. Some evaluative tests, however, are constructed somewhat differently. They are aimed at determining how many individuals, and the extent to which individuals, have attained a particular learning objective. These are called "criterion-referenced" tests.[19] In evaluation research it is often not as necessary to compare individual performance with the performance of other individuals as it is to compare individual performance with criteria which represent minimum levels of acceptable performance; i.e., a minimum level of acceptable performance is defined and set out as the criterion against which performance will be judged. "The tests are not meant to spread students out along a 'normal distribution' with a few doing very poorly, a few doing very well and the majority doing moderately well. Ideally, 100 percent of our stduents would obtain 100 percent of their objectives. In a norm-referenced test this indicates that the test was too easy; in a criterion-referenced test that the teaching and learning has been highly effective."[20]

Finally, it should be noted that, even if learning objectives have not been precisely specified and enunciated, the construction or selection of a test considerably defines the outcomes under investigation. Whether the test be norm-

referenced or criterion-referenced, its dimensions impose expectations and limitations on the behavior it elicits. Later we shall see that from one point of view this defining property of tests is a distinct advantage, whereas from another point of view it is a liability.

Before-and-After Design

In order to measure change which may be attributable to learning during an instructional program one requires a baseline; a measurement taken before the instructional experience begins and against which post-instructional behavior may be compared. Teachers are prone to making the assumption that prior to course attendance students have zero relevant knowledge. They sometimes make the error, therefore, of simply taking assessments after the instructional sequence and assuming that assessments which show highly acceptable performances demonstrate that teaching and learning has been effective. However, it should be clear that students do not always begin with zero relevant knowledge. Usually, some students have virtually all of the information or understanding necessary before the instructional process begins, some have parts of it, and a few have virtually none. A further alternative exists as well; some students may begin the instructional process having been badly misinformed and may have to "unlearn" a good deal before they can begin to deal appropriately with program content. It becomes obvious, then, that active assessment prior to the experience of the instruction is necessary if the instructor (or evaluator) wishes to know the status of

the learners at that time. If the same assessment is made after completion of the instructional program, comparison of the two sets of data should provide indications with regard to changes that may have taken place.

We have mentioned that the establishment of a "control" group is often not possible or feasible in such studies. The before-and-after design assumes that each subject is his own "control" and that pre-test measures are assessments of performance that would have occurred had there been no instructional intervention. Although this assumption may be questioned on a number of grounds, the before-and-after design--in some measure, does provide two sets of data which can be compared to provide indications as to the effectiveness of the instructional program.

Data Analysis

The emphasis on operational definition of objectives often means that information yielded by studies employing this paradigm is quantitative in nature. Much effort goes into developing objectively verifiable data-gathering methods such as counting of phenomena, events, and vital statistics. The acquisition of quantified information is, in fact, one of the major reasons for the definition of objectives in behavioral terms--so that measurements can be taken in so-called objective, numerical terms and subjected to statistical analyses. It is argued that the use of statistics in educational research is indicated for two primary reasons: (1) statistical techniques may be used to describe observations economically, i.e., they conserve the time and space necessary

to describe data; and (2) statistical techniques permit us to draw better inferences about the generalizability of observations.[21] If measurement can remain appropriate and still be reduced to counting procedures, then there is obvious value in quantification. One should not, however, assume that this automatically makes data more "objective" or that statistically-based data analysis techniques are always sufficient to meeting information needs. As Popham points out, "There is, unfortunately, no one-to-one relationship between the results of statistical operations and the judgements which should be rendered by the educational decision-maker. Statistics provide a tool whereby data can be parsimoniously described and more precisely analyzed than by merely 'inspecting the scores'. . . . But statistical results should not be equated with the final conclusions of scientific judgement. . . there is a crucial difference between a statistically significant result and a practically significant result."[22]

Objections to Definition of Objectives

A number of objections have been lodged against the classico-experimental paradigm, some more obvious, and some more serious, than others.[23] Much of the most serious criticism stems from the paradigm's central prescription to pre-specify behaviorally defined instructional objectives. Critics contend that such a statement of objectives is often a very long and tedious procedure requiring expenditures in time and effort which are of a very great magnitude.[24]

Parlett suggests four areas of concern here; (1) instructors very rarely even approach agreement on objectives; (2) the imprecision of language leaves objectives so generalized and diffuse as to be interpretable in numerous ways or so specific that literally hundreds must be listed to cover the full intent of even a few major course aims; (3) "getting an instructor to state his objectives requires his being highly articulate--which he may not be naturally--often about something he comprehends intuitively as sensible, but which he thinks would sound banal"; and (4) "Finally, there is the discomforting fact that actions speak louder than statements of intent".[25] Rowntree, though a proponent of the behavioral definition of objectives, admits that they can be exceedingly difficult to come by and dispels the myth that "science yields up non-trivial objectives more readily than the arts."[26] He also points out that there are no rules for adjudicating the merits and demerits of conflicting objectives; that the choice of objectives must ultimately be a political or social transaction, not a purely scientific undertaking.[27]

A related criticism is that concentration on the operational definition of objectives very often obscures questions about the worth of the objectives themselves and that operationally defined objectives are often a rather poor reflection of original intents; particularly if the evaluator(s) does not enter the drama until after the objectives have been set out.[28] In such cases one may find that the evaluation methodology can too easily concentrate on

improving the efficiency of teaching/learning programs that are basically manipulative and authoritarian in nature. In the same way, it is argued that since trivial and easily-measured objectives are the easiest to identify, they are often overemphasized.[29] Trow goes so far as to say, ". . . the most important and truly valued outcomes of higher education are extremely difficult, if not impossible to assess. As a result, many institutions, usually those with the least firm educational purposes and the least distinctive character, fall back in their self-assessments on those presumed outcomes of higher education that are more easily measurable."[30] Taking the argument into even more generalized climes, it is further suggested that our Western emphasis on well-defined rational, cognitive learning objectives is playing a central role in the accelerating development of "culture shock" and the initial inability to cope with ambiguous situations. If such is the case, there is need for educational experiences that ". . . have a more direct connection with life as it is lived in our relativist, kinetic, peripatetic, crisis-ridden society."[31] Be that as it may, even at the level of "simple" validation studies on instructional programs, there is a difficult problem in this area. It may be possible to develop relatively precise and easy measures of the immediate outcomes of a program; but it is often the long-range, difficult-to-measure outcomes which are most highly valued. Thus, we are faced with difficult-to-measure effects which are highly valued and more-easily-measured effects of much

lesser value. Clearly, emphasis on measurability can tend to produce emphasis on less valuable outcomes and inappropriately weighted research results.

A third, and rather serious criticism is that enquiries which are restricted to studying pre-specified outcomes seldom make a systematic search for other outcomes, thus, leaving unintended or unanticipated effects undetected. Intended or anticipated outcomes are often not the only important effects of an instructional program. An evaluation effort which does not remain flexible and open enough to detect other effects which may have a significant impact on people or systems is seriously circumscribed. The suggestion that prespecification of objectives and experimental design tend to create such limitations is labelled by Rowntree as a "baseless fear". He states that unintended results are no more likely to be overlooked than are the side-effects in drug trials.[32] This denial, however, is not backed by any evidence and the analogy turns out to be a poor one for his purposes. Overlooked side-effects resulting from limitations imposed by experimental methodology are rather well known in drug trials.[33]

To answer some of these criticisms of the definition of objectives, Rowntree says:

> Whatever you feel about the strength of objections to objectives, let's be quite clear about one thing. The denial of behavioral objectives does <u>not</u> mean that none are being achieved. The teacher who refuses to identify his objectives, or to admit that he has any, is nevertheless acting as an 'agent of change' on the behavior of his students. His every communication with students contributes to the achievement of some implicit objective. . . . The hidden curriculum marches on. If we

> don't specify the objectives we want, we'll have to put up with those we get."34

Although this would appear to be a valid argument for the specification of objectives in the curriculum design process, it does not necessarily have any application to the methodology chosen for evaluations of the program. The very fact that the hidden curriculum marches on is reason enough to criticize an approach which limits the inquiry to examining pre-specified and well-articulated objectives.

Criticism Regarding "Control"

We have already noted that in evaluation studies of instructional programs it is sometimes very difficult, perhaps impossible, to arrange for properly constituted control groups. It is often the case that there is no larger population from which a control group might be drawn. All prospective applicants for a program of instruction are actually admitted to the program, or those who are selected for inclusion display certain characteristics which make them significantly different from those who were not selected. Drawing a control sample from a population of somewhat similar individuals who did not apply for inclusion in the program is a doubtful procedure at best, even if control over a wide range of variables is practical and feasible. People who choose to enter a program are likely to be different from those who do not. Furthermore, assuming a situation with far more equally-well-qualified applicants than can be admitted, it is rare that individuals could be randomly assigned to "control" and "treatment"

groups; ethical and political considerations preclude it. So the number of instances where a well-chosen control group is even practically possible is rather small. Randomization and careful control may lend valuable strength to evaluation research, but it would appear that practical considerations often render them unusable, even counterproductive.

On a different, but related point, Parlett and Hamilton charge that the numerous relevant parameters which characterize an educational situation, if studied by use of the classico-experimental paradigm, create a requirement for very large samples. To maintain "tight control", each parameter must be "insulated" or "isolated" from others. If results are to be of value, relatively substantial samples must be used in studying each parameter. If the study is to offer more than an extremely limited investigation of one or two parameters, in isolation, large samples will be required. This, they say, runs counter to the widely acknowledged need for evaluation before large-scale application rather than after it. Further, attempts to limit sample size by imposing stringent controls can be not only ethically dubious but often administratively and personally inconvenient. "Even if a situation could be so unnervingly controlled, its artificiality would render the exercise irrelevant: rarely can 'tidy' results be generalized to an 'untidy' reality."[35] In much the same vein, but in a slightly different context, Gilbert, Light, and Mosteller--while reviewing methods for the evaluation of social programs--conclude that

although randomly controlled trials do have advantages,[36] three objections often emerge: "(1) randomly assigned treatments to sites or to people raises political problems. . .; (2) randomly controlled trials take too long; and (3) they are too expensive to conduct."[37] If all of these criticism are upheld, it would appear that Parlett is justified in his assertion that the more one thinks about it, the more this paradigm appears to be impractical.[38] Geismar, on the other hand, while recognizing that the design has its practical limitations as a measure of program outcome, reminds us that if it can be applied properly it offers the advantage of tightness of control--an advantage which is lacking in other designs.[39]

Other Limiting Properties

Other arguments suggest that adoption of the classico-experimental paradigm seriously limits an evaluation study in a number of other ways. In one case, it is argued that a before-and-after approach constrains the reseacher from adapting to changed circumstances. Contrary to assumptions underlying before-and-after methodology, instructional programs--especially highly innovatory programs--change a great deal during the period of study. "Because the prespecification of parameters, by definition, occurs at the outset, variables which emerge during the study are likely to be left out of the analysis. In an extreme case this neglect of 'new' variables may negate an entire evaluation study."[40] In like manner, it is argued that the design imposes artificial and arbitrary restrictions on the scope

of the study. By concentrating on the generation of quantitative data by objective means, the evaluator tends to neglect other data--regarding it as "anecdotal", "subjective", or "impressionistic". It is suggested that this neglected data may turn out to be highly relevant to the evaluation and that, in any event, the evaluator will probably be forced to utilize "subjective" information if he is to satisfactorily explain, weight, and contextualize his findings.[41]

In trying to allay the first of these objections, it is stressed that the initial statement of objectives need not be final and immutable;" . . . indeed it should make us more rather than less aware of the claims of the new objectives that arise urgently out of encounters with students in a learning situation. . . it should not deter any teacher from recognizing an unexpected educational opportunity and a new and valuable objective to which it relates."[42] Again, this counter-argument may have some weight when one considers the value of objective definition for curriculum design and implementation, but it may not apply to the evaluation procedure. Once objectives have been behaviorally defined, instruments designed, and pretests completed, there is little room for change. Borgatta and Jones suggest that such rigidity is not the fault of the design but a criticism of ". . . the status of the art and the underlying social sciences to which the applications refer."[43] Although many social science concepts and techniques are undoubtedly open to question, one cannot help but note that it is the paradigm itself which prescribes the

pre-specification of objectives and instruments, thus creating the basic rigidity.

With regard to the second criticism, Jones and Borgatta counter that: "Vague descriptions of activities as carried out are not the same as the statement of the consequences of these activities. . . . At some point, objectives should be stated in terms of measurable change in intended directions, and if reasonably identifiable criterion variables are not available, evaluative research is not feasible."[44] This pronouncement initially impresses as "down-to-earth", though somewhat moralistic; but it is based on a definition that mistakenly equates "evaluation" with "validation". As we noted in Chapter I, the validation of a program is not identical with the evaluation of a program. In any event, what are we to do if quantifiable criterion varibles are, for whatever reason, simply not available? Must we forgo evaluation entirely? It would appear that strict insistence on objective measures may lead to no evaluation at all.[45]

Insensitivity

A fourth area of concern is with the paradigm's insensitivity to any effects which do not follow the most apparent trends in the data and, additionally, with its failure to articulate with the varied questions and concerns of the numerous parties involved. Of the former criticism, Parlett and Hamilton say, "Atypical results are seldom studied in detail. Despite their significance for innovation, or possible importance to individuals and institutions

concerned, they are ironed out and lost to discussion."[46] This is a natural consequence of the emphasis on large samples and efforts to seek statistical generalizations. On the second point, "Since classical evaluators believe in an 'objective truth' equally relevant to all parties, their studies rarely acknowledge the diversity of questions posed by different interest-groups."[47] For example, the classico-experimental paradigm may prove useful for determining the dimensions of a program's overall immediate intended effects, but it does not go very far in explaining the causes of behavioral and social phenomena in the classroom; nor does it provide a constructive base for discussions concerning future instruction. Indeed, it does not even go very far in explaining reasons for the immediate outcomes--even if they are adequately measured.[48]

Outcomes Measurement

It would appear that the major effort in using the classico-experimental paradigm is expended toward measurement of outcomes which derive from pre-specified instructional objectives. A final line of criticism suggests that not only are other elements within the instructional system worthy of study (inputs and processes), but that the measuring techniques available for use within the paradigm are often inadequate to the task it sets for itself. Gilbert, Light, and Mosteller note that some effects may be too subtle to measure easily or inexpensively with instruments and procedures requiring heavy infusions of time, effort, and money.[49] We have already noted Trow's charge that the

really important instructional objectives in higher education are beyond such measurement techniques. In an effort to gauge the larger effects of an instructional program "outcomes" such as grades attained, drop-out rates, and achievement of scholarships are sometimes used. Trow allows that these sorts of phenomena are important in themselves--they are important determinants of the individual's future opportunities--but points out that they are inadequate measures of the outcomes of an educational experience. "For example, they are poor measures of the success of a liberal education in refining sensibilities, developing capacities for independent and critical thought, for the use of reason and evidence in everyday life, or for the enhancement of the individual's capacities for enjoying life and making fruitful contributions to it."[50]

As regards the study of other "worthy" elements in the instructional system, it is noted that concentration on outcomes appraisal (validation) very often leads to a complete neglect of other information needs. The classico-experimental approach tends to ignore examination of processes. As Weiss puts it, "The evaluation question as posed ignores the issues of why the program succeeds or fails. The <u>why</u> is often just as important as to know <u>how well</u> the program works."[51]

Proponents of the paradigm admit that some of this criticism is valid but counter with the observation that many available empirical testing techniques are entirely practical and feasible for evaluation studies of instructional

programs, and strongly re-assert that specification of operationally defined instructional objectives which guide the development and application of such techniques is equally realistic.

> Certainly we have a lot to learn about describing and exemplifying high-level and creative behaviors. Yet it is surely nonsense and flying in the face of a discipline's critical standards, to assert that there are no ways of stating in advance the kinds of quality one would look for (and the errors one would expect to see avoided) in a student's painting, essay, short story, musical score, research report or whatever. Professionals constantly make such appraisals and judgements about their colleagues' work, and examiners likewise (with even greater show of accuracy) for candidates in the arts. So what behaviors are they looking for? And when they are willing to externalize their criteria for judgement, we shall have our objectives.52

Socio-Anthropological Paradigm

In contrast to the major focus on outcomes stressed by the classico-experimental paradigm, the socio-anthropological approach concentrates primarily on processes occurring within the instructional system. As its name implies, this paradigm is heavily based on the investigative methods used by anthropological fieldworkers--methods which are concerned more with description and interpretation rather than measurement and prediction.[53] The anthropological fieldworker often lives with the social unit he is studying and conducts a wide-ranging investigation, using a variety of information-gathering techniques. In the initial stages his overall aim is to develop an understanding and portrait of the culture by building up a model which is constantly modified as new knowledge is acquired. As his understanding grows, he will be able to concentrate attention more

specifically on phenomena that have emerged in Stage 1. His end product will be an interpretation of the elements and relationships which compose the culture, and which is based on a distillation of the mass of information collected. A major emphasis in this approach is, therefore, on interpretation; ". . . on building up explanatory models of particular systems; on discovering patterns of coherance and interconnectedness. . . ."[54] Proponents of adopting the socio-anthropological paradigm for evaluation, or more correctly for "illumination"[55] of instructional programs, speak of its use in studying "innovatory" programs--"including 'evolutionary changes', 'experiments', and 'applications of educational technology'".[56] In short the adjective "innovatory" may apply to almost any instructional program since the actual implementation of any instructional sequence is seen as an innovation to one extent or another. To understand this viewpoint more completely it is necessary to recognize the crucial distinction between statements of the instructional system ("summaries") and implementations of the instructional system ("learning milieux"). The statement of instruction is the formalized plan or "script" which describes the instructional system. Sometimes it is set down in writing, more often it resides only within the instructor's mind. The classico-experimental evaluator most often begins by examining this plan and abstracting aims and objectives which are to be used as criteria against which the "goodness" of the final program will be measured. Although this procedure seems theoretically

proper--one should clearly state one's intentions and measure performance against their attainment--the emphasis upon pre-specified criteria ignores the fact that once adopted the plan almost immediately begins to go through constant modification as it is translated into teaching/learning behaviors. "The instructional system may remain as a shared idea, abstract model, slogan, or shorthand, but it assumes a different form in every situation. Its constituent elements are emphasized or de-emphasized, expanded or truncated, as teachers, administrators, technicians, and students interpret and re-interpret the instructional system for their particular setting. In practice, objectives are commonly re-ordered, re-defined, abandoned or forgotten. The original 'ideal' formulation ceases to be accurate, or indeed, of much relevance."[57]

The learning milieu is seen as the environment in which instruction occurs. It ". . . represents a network or nexus of cultural, social, institutional, and psychological variables. These interact in complicated ways to produce, in each class or course, a unique pattern of circumstances, pressures, customs, opinions, and work styles which suffuse the teaching and learning that occur there."[58] In short, the "learning milieu" is an environment or culture. Reasons for the choice of an anthropological research paradigm should now be obvious. It is an attempt to take into account contextual as well as experimental variables. The task does appear to be essentially similar to those undertaken in anthropological fieldwork. If we

accept the assertion that, "The learning milieu concept is necessary for analyzing the interdependence of learning and teaching, and for relating the organization and practices of instruction with the immediate and long-term responses of students",[59] then we will wish to select a research paradigm which will assist us in discovering and analyzing that learning milieu.

This is not to say that the discipline of social anthropology can offer us a standard methodological kit which will answer our multitudinous questions--merely that it offers a very general research strategy which may be adopted and adapted to the numerous and diverse learning cultures we wish to study. In so doing we are provided with a methodological prescription for investigating: basic values and attitudes; relevant personality characteristics; the dynamics of interpersonal and inter-organizational relationships; subtle learning-related phenomena such as independence of mind, judgement and creativity; pedagogical and andragogical procedures, forces and issues; teaching and learning styles; organizational history and/or a myriad of other phenomena, depending upon how they relate to the instructional program as a part of the learning milieu. Our aim will be description and interpretation, leading to explanation, and the resultant investigatory process may be characterized as having three parts: observation, further inquiry, and explanation.

Investigatory Steps

In the initial stage the investigator is primarily concerned with thoroughly familiarizing himself with the milieu. Since there will be no published accounts of this particular culture he must immerse himself in its day-to-day realities, combining personal, eyewitness observation with information obtained from informants' descriptions of behavior in order to produce an "holistic" portrait of the learning milieu. He will collect a very wide range of information so that he might begin to: isolate significant features for further study, discern patterns of correlation and causation, comprehend relationships between beliefs and practices, and discover connections between organizational patterns and individual behaviors. In time he will discover a number of common incidents, recurring trends, and frequently raised issues from which propositions about the culture (hypotheses) may be constructed.

In the second stage the investigator focuses his attention much more closely on the apparent significant regularities already discovered. Having become "knowledgeable" about the system, his inquiry now becomes more directed, systematic and selective--aimed at answering the most significant questions emerging from Stage 1 (testing hypotheses). He will attempt to gather information relating to the tentative propositions while, at the same time, attempting to remain as open as possible to new information which suggests alternative theoretical formulations.

In the final stage the investigator will attempt to weigh alternative interpretations of reality in the light of the information obtained. He tries to place regularities, congruities, and discongruities in a broader explanatory context. He tries to establish general principles which underly the functioning of the program and he outlines apparent patterns of cause and effect. In short, he attempts to distill sound theoretical formulations which explain the culture's most significant relationships, regularities and central issues. This process of gradually refining areas of concentration toward interpretation of significant patterns is illustrated by the "Integrated Studies Project".[60]

> Obviously the three stages overlap and functionally interrelate. The transition from stage to stage, as the investigation unfolds, occurs as problem areas become progressively clarified and re-defined. The course of the study cannot be charted in advance. Beginning with an extensive data base, the researchers systematically reduce the breadth of their enquiry to give more concentrated attention to the emerging issues. This 'progressive focusing' permits unique and unpredicted phenomena to be given due weight. It reduces the problem of data overload; and prevents the accumulation of a mass of unanalyzed material.61

Data Gathering

To accomplish this task the investigator takes a further leaf from the anthropologist's book; using a variety of media for gathering his information. He will almost certainly engage in direct observation of events in the instructional sequence, he may interview participants and examine documentary sources, he may use survey questionnaires,

tests and diaries. However, none of these media will be used in isolation; they will be used in tandem to view the issues or problems at hand from a number of different angles. He recognizes that each information-gathering procedure has its own limitations and attempts to cross-check and validate the tentative findings produced by each by "triangulating" on the objects of inquiry.[62] In Weiss' terms: "Adequate indicators of success in evaluation, like adequate measures of concepts in social research, usually entail multiple measurement. Each specific measure is an approximation of the outcome in which we are really interested."[63] If several plausable indicators of the same phenomenon concur as to its dimensions, effects, influences, etc., the validity of the observations is strengthened. If the several indices do not agree, the researcher has good reason to question their respective biases and reconsider his original conceptualization of the phenomenon.

Objections

It is not possible to draw sharp differentiations between the so-called "hard" and "soft" science approaches to investigation. Inasmuch as they attempt to accumulate systematic and reliable knowledge by empirical observation and the interrelating of concepts referable to empirical observations, both may clearly be designated as "science". The evaluator who adopts a socio-anthropological approach for his investigations may well be seen as ". . . actively engaged in accumulating raw data or in putting together generalizations based on observational data in a search for

systematized and reliable knowledge about human behavior."[64] Critics, however, suggest that the approach tends to encourage acceptance of generalizations on the basis of evidence that will not withstand close logical scrutiny, because of a failure to define concepts operationally. In the extreme, it is argued that in order to be meaningful a concept must be defined in terms of the measurement operations that would be involved in detecting it.[65] However, as Nettl points out, the prescription that all concepts employed must have direct and accessible empirical referents reduces the investigation to ". . . a niggling business, doing the easy thing because it is accurate and avoiding the difficult thing because it is imprecise."[66]

> The point of theorizing is to invent things that might be there, to pretend that they are there, because by imagining the existence of such things we can make better sense of whatever we feel we can observe in the real world. You simply cannot write a theory in sociology, or in any other discipline, without including unmeasurable, make-believe, or hypothetical notions.67

At the same time, one cannot adequately explain social phenomena by observation and interpretation of them ex post facto in a framework which simply feels right. For what feels "right" to one person may not feel at all "right" to another. Here the crucial and underlying issue of "subjectivity" becomes manifest. Apologists for the socio-anthropological paradigm do not deny an element of subjectivity in their work--it is inevitable, they say. They counter with the observation that no forms of research are free from subjectivity; all require skilled human judgements. Acceptance of this unextinguishable need for

subjective assessment leads them to encourage, rather than discourage, the use of interpretative human insight and skill. "The illuminative evaluator thus joins a diverse group of specialists (e.g., psychiatrists, social anthropologists and historians), where this is taken for granted. In each of these fields the research worker has to weigh and sift a complex array of human evidence and draw conclusions from it."[68] In order to reduce the likelihood of gross partiality on the part of the investigator it is suggested that a number of precautionary tactics be employed. For example:

> During the investigation different techniques can be used to cross-check the most important findings; open-ended material can be coded and checked by outside researchers; consultants to the evaluation can be charged with challenging preliminary interpretations and playing devil's advocate; and members of the research team can be commissioned to develop their own interpretations. At the report stage, in addition to the findings, critical research processes can also be documented theoretical principles and methodological ground rules can be discussed and made explicit; criteria for selecting or rejecting areas of investigation can be spelled out; and evidence can be presented in such a way that others can judge its quality.69

Bacon observed that ". . . as an uneven mirror distorts the rays of objects according to its own figure and section, so the mind, when it receives impressions of objects through the sense, cannot be trusted to report them truly, but in forming its notions mixes up its own nature with the nature of things."[70] Such observation cannot be totally circumvented--even by the most rigorously "objective" methodologies. In the final analysis, subjectivity remains. Even in the case of tangible phenomena

measured by external instrumentation, though the data may be highly reliable and valid, knowledge of them does not exist without human interpretation. Arguments about the pervasiveness of subjectivity do not, however, deny the merit of criticism regarding the innate subjectivity of the socio-anthropological approach. Operationalization does form a link between the mythical world of science and the "real world "; the lack of it destroys a link. Furthermore, though it may lend a measure of validity which would otherwise be absent, the Principle of Triangulation may also be seen as ". . . a cunning way of missing one bird with several stones."[71]

Another area of criticism focuses on the undeniable affects of the investigator and his activities on the phenomena he is investigating. When measuring intangibles the very act of measurement changes the value of the variable. Particularly when subjected to techniques such as direct observation and interview, the individual's behavior is likely to change in response to the evaluator's activity. Even without the investigator's intervention, his subjects are changing continuously and we do not have adequate tools for measuring dynamic systems--much less techniques that compensate for the influence of measurement upon them. Those espousing the socio-anthropological paradigm are unable to refute arguments of this sort. Again, they can but note that the same criticisms are also valid for other forms of data-collection on human subjects. Even a stringently controlled experimental study cannot entirely control for

such effects. For example, an experimenter may try to control for Hawthorne effect by comparing treatment and control group data, both of which are constant as far as the fact-of-being studied is concerned. However, it remains entirely possible that extraneous variables will interact with the Hawthrone effect constellation, or that the constellation will interact with experimental variables in unknown ways. Thus, they say, no design is impervious to such criticism. They suggest, however, that steps can be taken within a socio-anthropological study to minimize investigator-linked disturbances. "Illuminative evaluators recognize this and attempt to be unobtrusive without being secretive; to be supportive without being collusive; and to be non-doctrinaire without appearing unsympathetic."[72] This points to the conclusion that the socio-anthropological investigator requires much more than technical and intellectual capacity; he also needs a high level of interpersonal skill and personal integrity. Further, it means that from the outset the investigator must be very clear about his role and must communicate openly about the purposes of his study.

A further critical concern revolves around the paradigm's inductive nature and apparently limited scope. The paradigm may be of some use in studying small-scale programs but how possible is it to move from the particular to the universal? Does it allow for the sorts of generalizations we usually wish to make after expending so much time and effort on a research project? The apologist replies that instructional systems, despite their diversity, share many

characteristics. They are limited by similar conventions, divisions and problems. "There is a wide range of overlapping social and behavioral phenomena that accompany teaching, learning and innovating. This is widely acknowledged. However, few of these phenomena have been pinpointed, adequately described or defined accurately. Illuminative evaluation aims to contribute to this process. There is a need for abstracted summaries, for shared terminology, and for insightful concepts. These can serve as aids to communication and facilitate theory-building."[73] However, there is more to the criticism than is answered here, for any inquiry regarding generalizability is an inquiry about representativesness. How representative of the entire universe of students in an instructional program are the few who are interviewed or observed? How representative of the universe of opinion on a course is one chance remark overheard over coffee? The "illuminator" can only respond by trying to define his universe as extensively as is practical and by exposing his sampling methods to scrutiny; and again, he may argue that the same necessity attends any other approach to the collection of information.[74]

As to comments about the scope of the study--"A school that aims at too many targets may end up by hitting none"--it is argued that scope is a major strength of the socio-anthropological paradigm. It does not suffer from the severe limitations imposed by the pre-specification of objectives. Furthermore, just as anthropological fieldworkers do not enter the field with their minds completely blank,

the evaluator using a socio-anthropological approach specifies major areas for investigation in advance.[75] Remaining open to developments does not completely preclude initial conceptualization in broad terms.

A final area of criticism is concerned with the socio-anthropological paradigm's emphasis on description of process. Critics argue that it is all very well to describe ad infinitum but that inferences must be drawn at some point if the study is to be of any real value. The "illuminative" evaluator reminds critics that his aim is not simply descriptive but also explanatory. He notes that detailed descriptions of behavior using the anthropological approach tend to have high levels of "face validity" and "construct validity",[76] and demonstrates that the development of statements about patterns is not simply a descriptive account--". . . the pattern that emerges as the descriptive datum is in fact a generalization . . . it grants to its user some level of predictability based on a non-random occurrance of the particular phenomenon under observation."[77] The critic replies that this may be so but that the imputation of causes is lacking. We may gain knowledge of regularities and learn to predict them but we cannot say why they occur. The investigator responds that the imputation of causes may be at a low level but that it is certainly not totally lacking. Constant effort is made to discover threads which relate phenomena causally. This, he says, is more than can be said for most applications of the classico-experimental approach since they deal in

correlational relationships only. The critic may counter that talk about the discovery of patterns is misleading because, in the final analysis, the "patterns" are simply a product of the socio-anthropological investigator's emerging theory. He rejects any implication that such patterns somehow inhere in "deep structures" of reality.[78]

So the arguments continue--'round and 'round. One interesting point that should be noticed in considering this cacophony of criticisms and counter-criticism is that critical statements are usually phrased in "either-or" terms; either your sample is representative or not, either your data is subjective or it is objective. On the other hand, in each case the response is, "yes, we may be faulted to some extent but look at your own methods--yours are even worse". Although this counter-attacking response seems as adequate as is possible under the circumstances, the regularity of its form should make us pause for a moment. Whereas criticism has been levelled in "either-or" terms, rebuttal appears in a form which indicates that there are degrees of "goodness" or "badness" involved. It would appear that the criticism might be more forceful were it phrased in the same terms as the rebuttals, i.e., "the socio-anthropological paradigm espouses data gathering media which are more open to contamination by investigator bias than are certain other media"; "the paradigm rests less heavily on operationally defined referents and is, therefore, more open to diverse interpretations according to whim". Stated in this manner, the critical statements carry more

impact and render both sides of the controversy more open to close examination with regard to the practices which actually occur as the paradigms are adapted to use. Perhaps some comparison of the implied methodologies as they are put into practice can then be made. A beginning attempt to do just that will be our aim in Chapters VI and VII. Chapters IV and V present evaluation reports on two very similar instructional programs using the two different investigative approaches. Chapters VI and VII will attempt to contrast the methodological procedures implied by, and used in each.

CHAPTER IV

STUDY I

This Chapter comprises a report in which it is my intention to present information of possible use to the evaluation of a short sequence of instruction entitled "A Unitary Approach to Social Work Practice"; a workshop organized by the National Institute for Social Work (NISW), London, England, May 24 - 28, 1976. The sequence was arranged by the NISW in order to more thoroughly acquaint selected British social workers with "unitary" or "integrated" approaches to social work practice; approaches such as those promulgated by Pincus and Minahan, Goldstein, Whittaker, and Middleman and Goldberg.[1] The number of workshop participants was limited to twenty-eight, all of whom had substantial prior experience in social work practice and formal social work qualifications of one sort or another. Leadership was provided by four NISW tutors, each of whom had previous experience in leading short instructional programs on similar content. Of these, four, a designated senior tutor assumed primary responsibility for the organization and presentation of program components, but worked in very close co-operation with her three colleagues. An NISW "course organizer" undertook necessary administrative support tasks such as liason with the accommodating facility, participant orientation, registration, and correspondence. Her efficiency left workshop tutors relatively free to concentrate on the more direct instructional

tasks. Two other persons rounded out the workshop compliment: a tutor-in-training and myself, the course evaluator. The tutor-in-training attended instructional sessions and participated as if he were a registered workshop member. He did not assume any explicit instructional role but did attend and contribute to tutors' meetings. His primary interest appeared to be in observation of the workshop and interaction with participants, in order to further develop his own knowledge and skills relevant to future instructional activities.

The evaluation project was undertaken as part of a Ph.D. study which contrasted two investigative paradigms of possible use in evaluating instructional sequences for social workers. In accordance with the working agreement negotiated between the tutors and myself prior to the onset of instruction, I attempted to refrain from actual participation in the workshop sessions. My role might be best characterized as one of "non-participant observation"; I applied the necessary research instruments and observed many facets of the program in operation, but strove to minimize my own influence on the course of the instruction. The Study may be seen, therefore, as overwhelmingly "summative" in nature. It was intended to be a broad evaluation, presented <u>in toto</u> after workshop completion. The tutors had specifically requested that I refrain from providing any evaluative feedback during the course of the instructional sequence.

The workshop was held at the Anglian Regional Management Centre, Chelmsford, Essex. This is a purpose-built training centre which caters to a variety of short instructional programs and simultaneously maintains an in-house management training program. All instruction took place on the premises and almost all workshop participants, including the tutors, resided at the Centre for the duration of the workshop. In terms of both instructional and residential amenities, the Centre appeared to be entirely adequate to participants' needs and program requirements. Modern meeting rooms and residential accommodation, adequate meal-time and bar facilities, good food and service, and a wide range of audio-visual teaching aids were available.

Prior to attendance, the course organizer posted to each participant a kit of written material which included: a number of mimeographed papers dealing with elements of course content; a short list of suggested readings; and information on anticipated arrival times, arrangements for accommodation, tuition fees, travel arrangements, etcetera. The introductory papers summarized and briefly explicated the basic conceptual structures underlying unitary approaches to social work practice. They were closely grounded in material from the four textbooks noted earlier; the four major items on the suggested reading list. Tutors expected that very few participants would have the opportunity to read all four books prior to attending, thus, they strongly encouraged members to peruse the introductory papers.

Tutors viewed the instructional sequence as a participatory workshop within which their major role would be to assist and encourage members in developing a deeper understanding of unitary approaches to social work practice. Their chosen stated aims were threefold: (1) to help workshop members in exploring and understanding the basic concepts which make up a unitary perspective on social work practice, (2) to assist participants in learning to view the practice tasks associated with "assessment" from a unitary perspective, and (3) to encourage workshop members in learning to apply a unitary approach when "intervening" in problematic practice situations.

After an initial "Welcome and Introduction" session, the workshop membership was divided into four small work groups and a tutor was assigned to each subgroup.[2] Except when the entire membership met in plenary sessions, participants spent all scheduled time on Monday, Tuesday and Wednesday working on assigned tasks within these small groups. Monday afternoon and evening were programmed for small group discussions related to an assigned hypothetical case study. Tuesday morning was taken up with a plenary session, half of which was concerned with subgroup reporting on the previous day's work and the other half with a presentation of various unitary approach models by the senior tutor. Tuesday early afternoon was scheduled as "free time", while from 4:30 P.M. until after 9 P.M. participants engaged in small group explorations of "systems" concepts. These deliberations were reported upon in a

short plenary session. Wednesday followed much the same pattern; small group discussion of a second hypothetical case study, followed by a pre-dinner plenary session to discuss progress within the subgroups and a programmed re-organization of subgroups for the next day. Wednesday evening was left free.

Four new subgroups were constituted for Thursday's sessions. Participants were encouraged to form small groupings which reflected loci of specific interests and individual preferences with regard to subject-matter. Each group, therefore, undertook a somewhat different discussion topic, but each topic was associated with the application of a unitary approach to social work pracrice in a particular setting. Except for a free period in the early afternoon, small group discussions continued until after 9 P.M.

In the final plenary session on Friday morning, achievements of the subgroups and conclusions arising from their activities were the focus of early attention. This discussion gradually developed into an examination of major workshop themes and issues, culminating in an evaluation session within which all participants and tutors contributed to an orally-presented critical review of the entire instructional sequence.

Thus, from Monday afternoon until Friday noon, participants were engaged in a minimum of seventeen hours of small group activity and nine hours of plenary discussions. Further, the residential nature of the course encouraged much informal participant interaction directly concerned

with workshop content. It was apparent, for example, that a high proportion of scheduled "free time" was spent in informal discussions of unitary approaches and their applications to practice.

Instrumentation

We noted above that the workshop concluded with a plenary session which was partially devoted to eliciting evaluative feedback. As course evaluator, I attempted to systematically gather additional information of probable use to evaluative decision-making by administering four paper-and-pencil instruments: (1) a "Personal Information" questionnaire, (2) a performance test based upon two hypothetical case studies, (3) a set of "Daily Evaluation" rating scales, and (4) an "Overall Evaluation" rating form. Copies of these instruments may be found in the Appendix.

The "Personal Information" form was a straightforward questionnaire, administered to all workshop members prior to attendance and requesting: basic personal information such as age, sex and mailing address; information on employment experience and educational qualifications attained; and information on the amount of preparatory reading completed prior to instructional onset.

In order to assess, in a relatively direct manner, the extent to which the workshop actually met tutors' primary stated instructional objectives, a performance test was devised and administered on two occasions; once prior to the workshop experience and again upon workshop

completion. Two similar hypothetical case studies were constructed from details of actual practice situations; details of which participants could have had no prior knowledge. All subjects were requested to respond to these case studies by producing a set of written "notes" which might be used in a hypothetical discussion of the case material. Prior to the workshop, one half of the prospective participants--chosen at random by means of a table of random numbers--were asked to prepare a set of "notes" on Case A, while the other half were asked to respond to Case B in similar manner. Upon workshop completion, all participants received the case study to which they had not responded in the first instance, i.e., the original Case A group received Case B material and vice versa. This procedure eventually resulted in thirty-two complete sets of "notes", two from each of the sixteen respondents who carried out both segments of the performance test and submitted their responses for analysis. All thirty-two sets of "notes" were then rewritten to conform to a standardized structure and were presented to a panel of four NISW tutors for judging. Although this group of tutors was not identical with the group who led the workshop, each was experienced in leading instructional sequences on "unitary approach" content. Judges were given standardized sets of "notes" in order to minimize the possibility of their distinguishing between individuals, or between pretest and posttest responses from the same individual. Their task was to rate each set of "notes" on

three dimensions: demonstrated understanding of a unitary perspective on social work practice, demonstrated ability to use a unitary perspective in assessing the case situations, and demonstrated ability to use a unitary approach in developing interventive strategies for application in the case situation. Each dimension was to receive a rating on a eight-point scale which was anchored by "very poor" at the low end and "very good or excellent" at the high end. Simply, the general idea of the performance test was to compare each respondent's pretest scores with his own posttest scores, in order to determine whether any apparent changes had occurred as a result of the instructional sequence. As we shall see later, data were also analyzed in other ways.

At the end of each workshop day, participants were asked to complete a "Daily Evaluation" form which contained nine rating items. Items numbered 1 - 6 asked subjects to rate the day's sessions on six-point scales; poles were labeled "poor" (low) and "excellent" (high). Item stems dealt with: (1) "holding your interest"; (2) "amount you learned"; (3) "relevance to your needs"; (4) "amount and quality of work completed in your group"; (5) "organization, sequencing and timing of sessions"; and (6) "adequacy of tutors". Item 7 asked about the extent to which the usual participation pattern in the group was perceived to be "balanced". Item 8 dealt with the extent to which group members were perceived to have concentrated their collective attention on the assigned discussion topics. Item 9 asked

how well the group's tutor had controlled discussions. The five-point scale for this last item ranged from "should have controlled a great deal more", through "controlled just about right", to "should have controlled a lot less". Because of its rather different nature, this item was treated somewhat differently from the rest. It was anticipated that the combination of rating responses would produce a composite nine-dimensional picture of each day's small group activity, from the participants' points of view.

Once the workshop was completed, but before members left the facility, they were asked to undertake a more comprehensive "Overall Evaluation" rating form. Some of the items included were very similar to those on the "Daily Evaluation" forms, others were quite different. Twenty-six such forms were duly completed and returned. Approximately one month after the workshop, the same instrument was posted to all participants in an effort to gain initial follow-up rating data. Nineteen of these forms were completed and returned. To acquire more extended follow-up ratings, the same form was posted out a third time five months after the workshop. On this final application, twelve forms were completed and returned.

Description of Participants

Of the 28 workshop members, 15 were female and 13 male. Their mean age was 37.8 years and there was no statistically significant difference in age between females (37.1) and males (38.5). Based upon information gathered

on job titles and the percentages of working time spent in various activities: 13 reported themselves to be social work teachers, 4 were social work practitioners, 3 were senior social workers (supervisors), another 3 were training officers, 2 were teacher/administrators, 1 was a probation officer, 1 a fieldwork instructor, and 1 a senior administrator in a social service agency. All were, therefore, currently employed in the social welfare field and appeared to have been so employed for some time; their mean length of reported social work experience was 12.4 years. Each participant held at least one formally recognized educational qualification relevant to the practice of social work; in fact, they displayed an average of over two such formal qualifications per person. Of the four textbooks suggested in the preliminary orientation kit, only a few participants reported having read none, none had read more than three, and the mean reported amount of reading was 0.68 books per person.

The largest identifiable subgroup, the social work teachers, was composed of 7 males and 6 females. Their mean age was 37.7 years and their reported mean length of experience in social work was 12.8 years. They reported an average of 2.15 formal relevant qualifications per individual. On none of these characteristics did they appear to differ significantly from the reported characteristics of the workshop membership as a whole. With regard to reading, however, they were significantly different; as a group, they had read significantly more than the workshop average ($\bar{X}$ = 0.923,

$t = 1.875$, $df = 26$, $p < .1$). Since social work teachers are often expected to read more of the social work literature than agency-based personnel, this finding is hardly surprising.

We noted that the workshop membership was divided into four small discussion groups for the Monday, Tuesday and Wednesday sessions. The compositions of those four groups are represented in Figure 4-1.

- On average, Group A was composed of individuals who were slightly younger and less experienced in social work than the workshop membership as a whole. Members of this group also reported having done less than the average amount of preparatory reading; a difference, however, which did not show statistical significance on a t-test.

- The males in Group B were somewhat younger than the overall workshop average, and Group B members' average length of experience in social work was correspondingly lower than the workshop mean.

- The members of Group C were slightly older than the workshop mean. They also reported a longer mean length of social work experience and more formal relevant qualifications. More strikingly, Group C members' reported mean amount of preparatory reading appeared to be substantially higher than the overall workshop average; primarily because social work teachers were disproportionately represented in this group. The t-test did not, however, render a statistically significant result.

FIGURE 4-1

COMPOSITION OF DISCUSSION GROUPS

	Group A	Group B	Group C	Group D
Males	3	4	3	3
Females	4	3	4	4
Mean Age	34.6	36.4	39.6	40.6
Mean Age (Males)	36.0	35.3	39.0	45.0
Mean Age (Females)	33.7	38.0	40.0	37.3
Employment	3 teachers 1 prob. officer 1 practitioner 1 sen. soc. wkr. 1 train. officer	2 teachers 2 teach/admin. 1 practitioner 1 sen. soc. wkr. 1 train. officer	5 teachers 1 train. officer	3 teachers 2 practitioners 1 sen. soc. worker 1 agency admin.
Mean Length of Experience	9.57	10.57	16.09	13.21
Mean number of Qualifications	2.0	1.9	2.7	2.0
Mean Number of Texts Read	0.54	0.68	1.04	0.67

- While the females in Group D were of average age, the males were considerably older. As one might expect, members of this group reported somewhat longer social work experience than the overall workshop average.

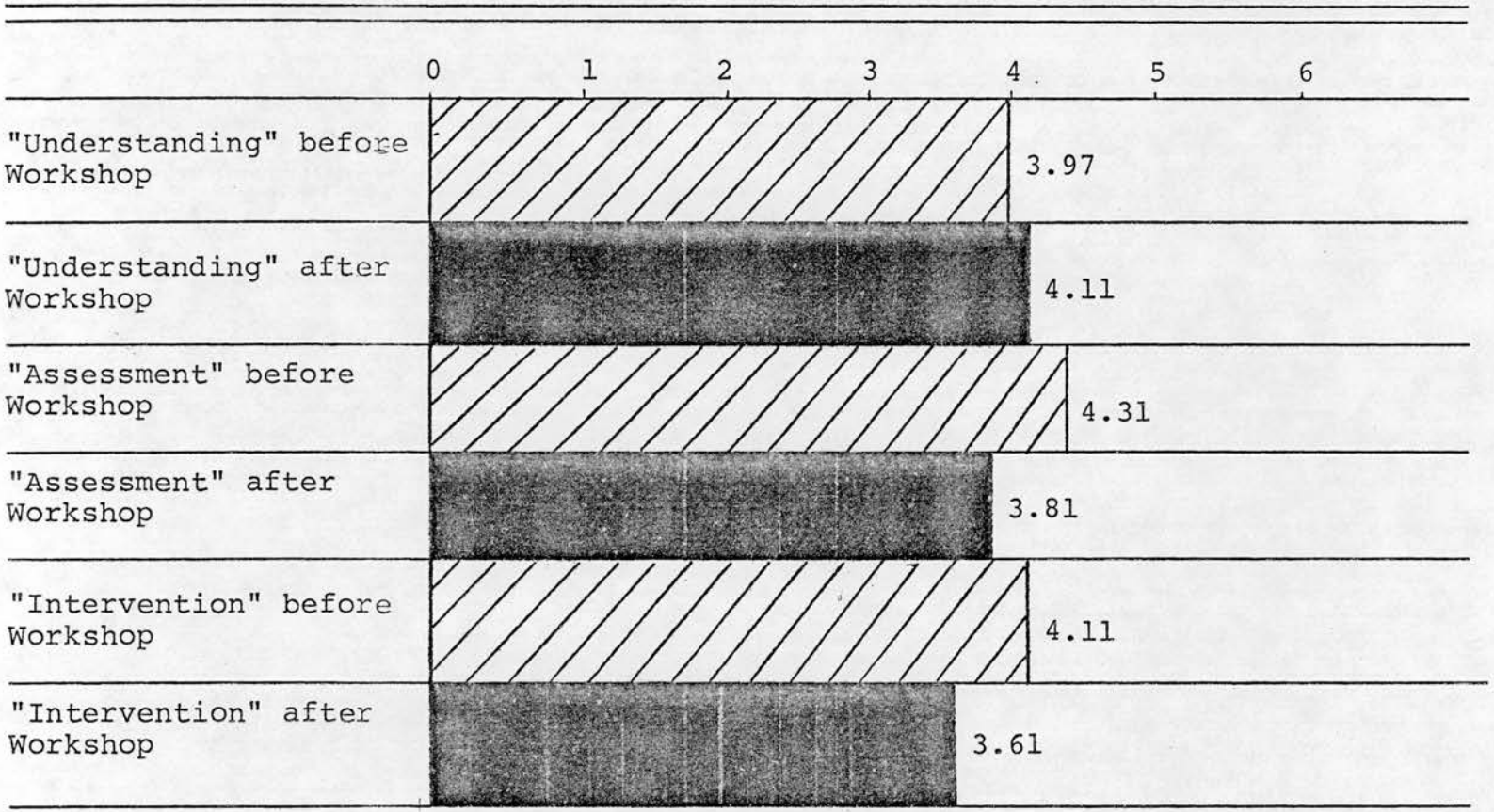
Results of Performance Tests

The overall results of the performance test (case studies) are graphically presented in Figure 4-2. They indicate a very slight increase on the "understanding" dimension and slightly larger decreases on the "assessment" and "intervention"dimensions. On closer analysis, none of these differences over time reached a level of statistical significance. Furthermore, none of the differences between the three test sections were statistically significant. So, although the overall difference between pretest and posttest performance was in a negative direction (i.e., subjects appear to have lost in ability as a result of the workshop), the fact that it did not reach proportions representing a statistically significant level suggests that it is the product of chance variations and not attributable to some sort of iatrogenic dysfunction. At the same time, it provides no evidence to suggest that workshop participants actually learned to understand, assess or intervene better as a result of their workshop experience.

Performance test results were also analyzed for differences between "agency-based" and "school-based" participants, and for differences between male and female subjects.[3] Although the agency-based group demonstrated an

FIGURE 4-2

OVERALL RESULTS OF PERFORMANCE TESTS



overall slight improvement in performance between pretest and posttest and the school-based group showed a slight overall decline, the difference between them was not statistically significant. Neither was there a significant difference between groups divided by gender.

The same data were also analyzed for differences between the initially formed working subgroups (Groups A, B, C and D) and one statistically significant performance difference was discovered. On the whole, members of Group D appear to have benefited from the workshop experience significantly more than members of Group A ($t = 2.27$, $df = 5$, $p < .05$). When compared with all other groups at once, however, no one group performed significantly differently from the others.

An analysis of differences between those subjects who showed the greatest positive change in performance and those who showed the greatest negative change produced only one tentative finding. Those subjects who showed the greatest amount of change in the negative direction appeared to be more likely to have attained "Home Office Certificates" as part of their formal relevant qualifications. The statistical difference between the group possessing Home Office Certificates and those who did not approached the .1 level of significance ($t = 1.47$, $df = 14$, $p > .05$); thus, this is not a clearly cut difference. The most we might conclude is that those possessing Home Office Certificates tended to show slightly more negative change in performance than other subjects.

An attempt was made to statistically correlate an "amount learned" index, obtained from the combination of five items on the "Overall Evaluation" instrument, with the amounts of change registered by the performance test.[4] No significant correlation was evident. Those who showed greater than average amounts of positive change in performance did not offer higher, or lower, ratings on the "amount learned" index. Similarly, those subjects who showed higher than average negative change in performance did not provide lower, or higher, ratings on the "amount learned" index. Further, no significant correlations between performance and other characteristics were noted. Correlations were attempted with: age of subject, length of social work experience, changes in "Overall Evaluation" ratings between administrations, an index of participant satisfaction with each day (combination of "Daily Evaluation" items numbered 1, 2, 3 and 4), overall ratings of the tutors (combination of "Overall Evaluation" item number 6 and all "Daily Evaluation" items numbered 6 and 9), "work on topic" ratings (all "Daily Evaluation" items numbered 9), overall "satisfaction with group" ratings (all "Daily Evaluation" items numbered 4, 7, 8 and 9), and the reported amounts of preparatory reading completed. Neither did performance correlate significantly with "Overall Evaluation" Administration I ratings (all items).

Reliability of Judges' Ratings

Performance test results were, of course, based upon ratings provided by the four NISW tutors who made up

the judging panel. A statistical analysis of interjudge reliability was undertaken by correlating the ratings given by each judge with the ratings given by each of the other judges. Then a reliability coefficient was obtained by averaging each of the three resulting correlation coefficients. In the event, interjudge reliability was rather poor. The only statistically significant correlation was between Judges B and D ($r = .638$, $p < .1$). Figure 4-3 graphically represents the extent of interjudge reliability by showing the obtained mean correlation coefficient for each judge.

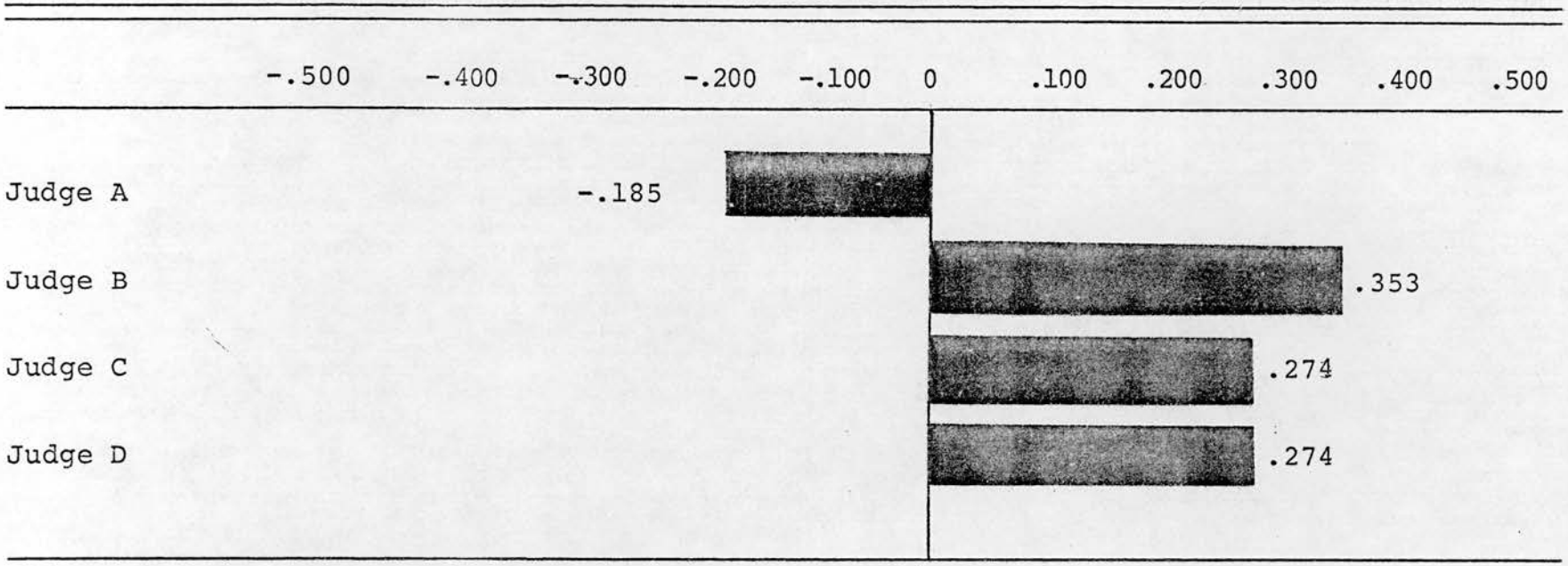
Obviously, no judge's ratings reflected a very high overall reliability coefficient. The ratings of Judge A negatively correlated with the others, but results were far from statistically significant levels. The ratings of Judges B and D did show a relatively high positive correlation with each other (see above), while those of B and C ($r = .486$) and C and D ($r = .506$) were somewhat lower.

Results of "Daily Evaluation" Ratings

The "Daily Evaluation" rating instruments completed by almost all participants on Monday, Tuesday, Wednesday and Thursday, were analyzed with regard to all items. On the "tutor control" items, however, only a partial analysis was possible because subjects found the item inappropriate to Wednesday's sessions and did not provide ratings for that day. Thursday definitely received the highest combined daily ratings. In fact, the difference between the combined

FIGURE 4-3

JUDGES RELIABILITY COEFFICIENTS



ratings for Thursday and the combined ratings for all other days was statistically significant ($t = 2.12$, $df = 30$, $p < .05$). The day which received the lowest overall ratings was definitely Wednesday; here the difference from overall ratings for all other days was even more pronounced ($t = 3.05$, $df = 30$, $p < .01$).

On an item-by-item analysis, participants apparently felt that Thursday most held their interest ($\bar{X} = .813$), provided them with the greatest amount of learning ($\bar{X} = .673$), was most relevant to their working needs ($\bar{X} = .793$), was the day on which the greatest quantity and quality of work was completed ($\bar{X} = .712$), was the best day in terms of orgnization and sequencing ($\bar{X} = .750$), was the day on which the tutors performed best ($X = .736$), was the day on which discussions were most "on topic" ($\bar{X} = .865$), and - of Monday, Tuesday and Thursday - was the day on which tutors showed the most balanced control over group deliberations ($\bar{X} = .865$). On the same characteristics, Wednesday received the lowest ratings: "holding your interest" ($\bar{X} = .600$), "amount of learning" ($\bar{X} = .480$), "relevance to needs" ($\bar{X} = .447$), "work completed" ($\bar{X} = .472$), "adequacy of tutors" ($\bar{X} = .646$), "organization and sequencing" ($\bar{X} = .550$), "on topic" ($\bar{X} = .780$), "balance in tutor control" ($\bar{X} = .720$). The only item which departed from this very strong trend was Item 7 ("balance in participation pattern"). On this item, Tuesday received the highest overall ratings ($\bar{X} = .864$) while Monday received the lowest overall ratings ($\bar{X} = .764$).

"Daily Evaluation" data were analyzed on a group-by-group basis also; the overall results are reflected in Figure 4-4. On Thursday members were not, of course, in their original groupings. The groups have been maintained statistically and graphically in order to indicate how the ratings changed with the changes in group composition. Participants from Groups A, B and C appeared to be generally pleased with the changes on Thursday; Group D members did not appear to be as pleased. Overall, Group D received the highest ratings and, though not significantly different from those of B and C, they were significantly higher than those of Group A ($t = 2.589$, $df = 16$, $p < .02$). Group A ratings were not significantly different from those for B and C.

The following results arise from analysis of the Monday, Tuesday and Wednesday ratings only.

- On Item 1 ("holding your interest") Group D received the highest rating ($\bar{X} = .74$) but was only slightly higher than Group C ($\bar{X} = .72$). Group A received the lowest rating ($\bar{X} = .59$).

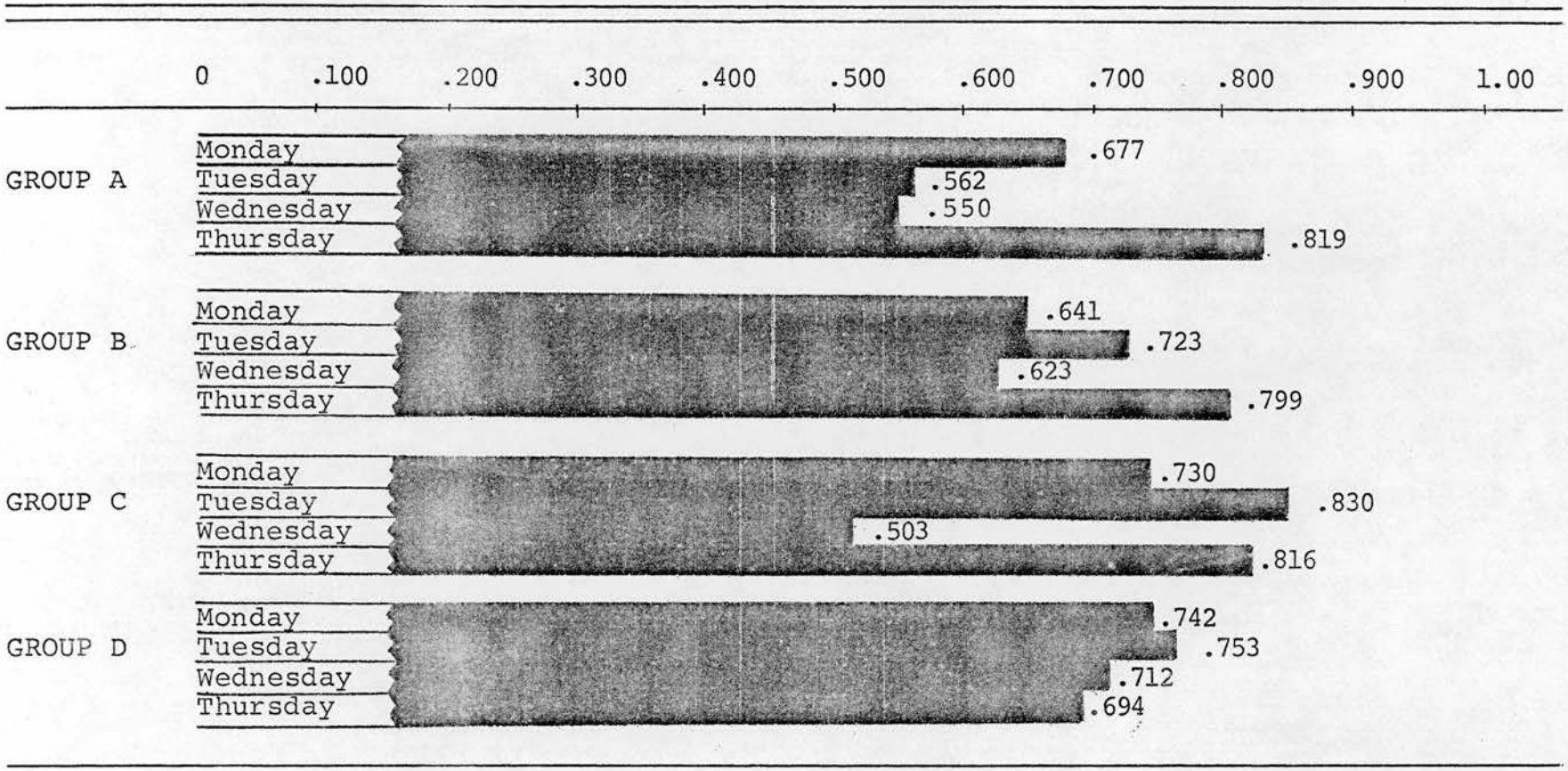
- On Item 2 ("amount you learned") Group D was given the highest rating ($\bar{X} = .627$), followed closely by Group B ($\bar{X} = .578$). Group A received the lowest rating ($\bar{X} = .420$).

- On Item 3 ("relevance to your needs") Group D was rated highest ($\bar{X} = .623$), while Group A was rated lowest ($\bar{X} = .407$).

- On Item 4 ("work completed") Group C and Group D received equally high ratings ($\bar{X} = .667$), while Group A

FIGURE 4-4

OVERALL "DAILY EVALUATION" RATINGS BY GROUP



received the lowest rating ($\overline{X}$ = .510).

- On Item 5 ("organization, sequencing and timing") Group D was given the highest ratings ($\overline{X}$ = .717). The other Groups received substantially lower ratings ($A\overline{X}$ = .587), $B\overline{X}$ = .550, $C\overline{X}$ = .583).

- On Item 6 ("adequacy of tutors") Group D received the highest rating ($\overline{X}$ = .703), followed closely by Group B ($\overline{X}$ = .687). Group A received the lowest rating ($\overline{X}$ = .603), but was only slightly lower than Group C ($\overline{X}$ = .613).

- On Item 7 ("balance of participation pattern") all Groups were rated similarly ($\overline{X}$ = approximately .830).

- On Item 8 ("on topic") Groups C and D received equal highest ratings ($\overline{X}$ = .833) while Group A was rated lowest ($\overline{X}$ = .707).

- On Item 9 ("balance in tutor control") Group D received the highest ratings ($\overline{X}$ = .855) and Group A the lowest ($\overline{X}$ = .695).

"Daily Evaluation" data were also analyzed for differences between agency-based and school-based subjects. Although the overall results tend to suggest that the school-based group were somewhat more favourably inclined to Tuesday than any other day, no statistically significant differences were found.

To summarize the "Daily Evaluation" ratings: Thursday received a significantly higher overall rating than all other days, while Wednesday was rated significantly lower than all other days; Group D received the highest overall ratings and there was a statistically significant

difference between the ratings for Group D and Group A; no statistically significant differences were found between ratings provided by school-based and agency-based workshop members.

Results of "Overall Evaluation" Ratings

The "Overall Evaluation" instrument, administered on three occasions, included the following rating items.

Item 1 - "Did you enjoy attending this workshop?"

Item 2 - "How much did you learn from the workshop?"

Item 3 - "How relevant was the workshop content to your on-the-job needs?"

Item 4 - "How well balanced was the workshop?"

Item 5 - "If this workshop was offered again, would you recommend that your colleagues attend?"

Item 6 - "How adequate were the tutors' provisions for student feedback and evaluation?"

Item 7 - "How would you rate the workshop's facilities as regards: (a) meeting space, (b) accommodation, (c) food and service, (d) accessibility, and (e) leisure opportunities?"

Item 8 - "To what extent do you think the workshop has:

(a) helped you to understand the unitary approach?

(b) improved your ability to apply the unitary approach?

(c) provided you with new ideas for agency and service reorganization?

(d) provided you with new approaches to assessment and intervention?"

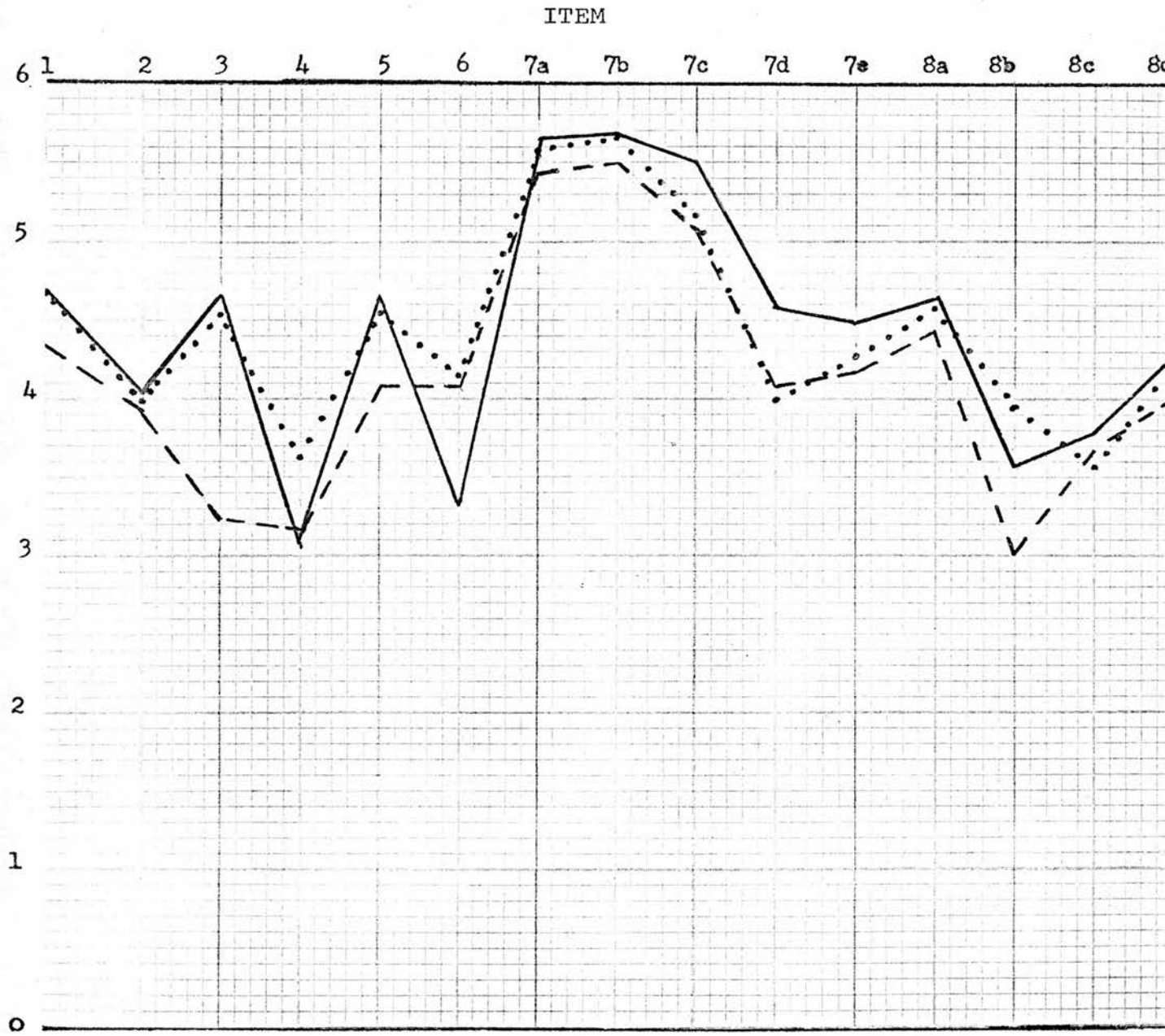
Subjects were requested to rate each item on a scale which ranged from 0 to 6 and had appropriate anchoring phrases at either end. The higher the rating, the greater positive regard indicated. Results of the three applications are presented graphically in Figure 4-5.

It would appear from the graph that all three sets of data positively correlate with each other. In fact, Application I and Application II (one month apart) show a rather strong positive correlation ($r = .911$, $df = 13$, $p < .05$). Although there appeared to be a very slight general decline in ratings over the intervening month, it was found to be statistically insignificant. Applications II and III (four months apart) also show a significant positive correlation ($r = .775$, $df = 13$, $p < .05$). Between these two applications there appeared to be a somewhat larger decline in rating levels but, again, it did not reach statistically significant proportions. The ratings taken five months apart, Applications I and III, also show a significant positive correlation ($r = .649$, $df = 13$, $p < .05$) but here the cumulative difference between Applications is statistically significant ($t = 2.258$, $df = 28$, $p < .05$). Thus, we have a relatively strong indication that the general decline between Application I and III reflects a real decline in the participants' overall regard for the instructional sequence. An item-by-item analysis produced the following findings.

- Item I ("enjoyment") was included primarily as a "starter"; it showed no significant change between

FIGURE 4-5

RESULTS OF "OVERALL EVALUATION" ADMINISTRATIONS I, II AND III



Administration I _____

Administration II . . . . . .

Administration III _ _ _ _ _ _ _ _

Applications I and III.

- Item 2 ("amount learned") showed a slight decline between Applications II and III. There was no significant correlation between Applications I and III and, although the decline between Applications I and III would appear to be fairly large, it does not reach the .05 level of statistical significance ($t = 1.862$, $df - 20$, $p < .1$).

- Item 3 ("relevance to needs") showed a slight decline between Applications I and II and a very distinct decline between Applications II and III. Applications I and III were positively correlated ($r = .639$, $df = 9$, $p < .05$) and the decline between them was also statistically significant ($t = 2.560$, $df = 20$, $p < .05$). Immediately after the workshop, participants appear to have considered the experience substantially relevant to their on-the-job needs; after five months experience on the job, they considered it significantly less relevant.

- On Item 4 ("balance") there was a considerable increase in rating level between Applications I and II but it disappeared almost entirely on Application III. No statistically significant differences were found.

- On Item 5 ("recommend colleague attendance") there was a slight decline between Applications I and II and a much more substantial decline between Applications II and III ($r = .518$, $df = 9$, $p > .05$; $t = 2.176$, $df = 20$, $p < .05$). Immediately after the workshop, subjects were quite positive about recommending that their colleagues attend; five months later they remained relatively positive

but the rating level had dropped significantly.

- Item 6 ("tutors' provision for feedback and evaluation") was the only Item showing a possibly significant increase over time. The increase between Applications I and II was even greater than the increase between Applications II and III. Applications I and III were positively correlated ($r = .703$, $df = 8$, $p < .05$) but the increase between them did not reach a level of statistical significance ($t = 1.330$, $df = 20$, $p > .05$). It might be hypothesized that this reversal of the overall trend toward decline over time was due to subject confusion with regard to my role as evaluator. The NISW tutors carried out no follow-up evaluation but it may be that my continuing evaluative activity was somehow perceived to be linked with tutor activity.

- Items 7 a, b, c, d, and e ("facilities") showed a slight overall decline between Applications I and III but it was not statistically significant.

- On Item 8a ("understanding") the same slight decline between Applications I and III was apparent. There was no significant correlation between Applications I and III.

- On Item 8b ("application") Applications I and III did show a significant positive correlation ($r = .811$, $df = 8$, $p < .05$) and the slight decline was not statistically significant.

- On Item 8c ("reorganization ideas") there was very little change in ratings between Applications I and III.

- On Item 8d ("approaches to assessment and intervention") Applications I and III showed a significant positive correlation ($r = .709$, $df = 9$, $p < .05$) and the decline in rating levels between them was statistically insignificant.

- On the learning index, obtained from the combination of Items 2, 8a, 8c and 8d, there was no statistically significant change between Applications I and III.

Conclusion

With regard to the tutors' stated workshop aims, to the extent that they may be interpreted simply in terms of providing opportunities for participants to examine and begin to apply the basic concepts of a unitary approach, the instructional sequence would appear to have been successful. Over the four-day workshop period, participants spent a minimum of twenty-five hours engaged in formal examination of essential unitary approach concepts and some of the possible areas of application in social work practice. In addition, the residential nature of the workshop provided members with a great deal of further opportunity for relevant, albeit less formally organized, discussion. Workshop members, therefore, appear to have spent a substantial proportion of the week exploring unitary approach concepts and discussing possible applications amongst themselves. Furthermore, a very large majority of evaluative ratings (both "Daily" and "Overall") were extremely positive - well above the scales' medians. This strongly suggests that the workshop

participants, themselves, believed the program to be a success.

If, however, one breaks down the tutors' stated aims into three more specific objectives (i.e., to promote the understanding of unitary approaches, to promote the acquisition of appropriate assessment abilities, and to promote the acquisition of appropriate interventive abilities), the limited evidence on subjects' performance change is not at all suggestive of success. The overall changes in performance between pretest and posttest were very small; none were statistically significant. One is likely to conclude, then, that no overall change, other than that produced by chance variations, took place. The possible presence of changes in the negative direction might suggest that other, unknown, factors were operating. One tends to expect some indication of change in a positive direction, even if only as a result of increasing subject sophistication arising from repeated test administration. If changes in the negative direction had been more marked one might have suspected that either the instrument measured some other phenomena which obscured any learning that occurred, or that some sort of iatrogenic dysfunction had taken place. Alternatively, one might have hypothesized that, for one reason or another, subjects tried harder on the pretest than they did on the posttest. Since the results were not statistically significant, however, one can but conclude either that the workshop experience did not affect any change in the performance characteristics measured, or that the instruments were insensitive to any relevant changes

which did occur.

With regard to characteristics of the participants and their workshop groupings, there were few statistically significant correlations with the amount of learning indicated by pretest - posttest performance differences. There was an indication that members of Group D had learned more than members of Group A; at least, Group D members demonstrated significantly less performance change in the negative direction. This finding tends to correspond with Group D's significantly higher "Daily Evaluation" ratings, but no clear-cut relationship was established. There also appeared to be a tendency on the part of subjects who displayed the greatest amounts of negative change in performance (i.e., appeared to lose most in "understanding" and the two "abilities") to have obtained "Home Office Certificates" as part of their formal qualifications for social work. One could hypothesize about various intermediary variables which might have intervened to produce this result, but the finding was not statistically strong enough to substantially rule out the probability that it might be accounted for simply by the affects of random variations.

With regard to the members' nine "Daily Evaluation" ratings of workshop elements, there is extremely strong evidence indicating that Thursday was generally perceived to be the best overall day, and even stronger evidence suggesting that Wednesday was perceived to be the worst overall day. The small group discussions on Thursday were held to be: the most interesting, those which encouraged

the greatest amount of learning, the most relevant to participants' on-the-job needs, the groups in which the greatest quality and quantity of work was completed, the best in terms of organization and sequencing, the day on which tutors' performance was at its best, the day on which groups concentrated most attentively on the topics at hand, and the day on which tutors showed the most balanced control over group discussion. On the same characteristics, Wednesday received the lowest ratings. Since Thursday's program was devoted to the application of unitary approach concepts in participants' working circumstances and since it was the day that the original small groups were disbanded, one might hypothesize that either or both of these factors represented valuable directional changes in the workshop's program. Although conclusions must be highly tentative, one might suggest that workshop members found special utility in discussing self-chosen topics with like-minded colleagues. One might also tentatively conclude that Wednesday's sessions might have profitably concentrated on more specifically job-related "application" topics, rather than on a second hypothetical case study.

As a marked departure from the general trend, Tuesday received the highest ratings for "balance in participation pattern". Since Monday was the first workshop day, one might have anticipated its low "balance in participation" ratings; group members were, after all, just settling in. At the same time, one might have expected Wednesday to capture the highest honours since group members were likely to have

developed more mutually satisfying participation patterns by that time. The results suggest, in fact, that by Wednesday the initial group configurations were fast becoming perceived as less useful. This finding, therefore, lends weight to a proposal that initial workshop groups be dissolved earlier; perhaps on Wednesday morning. Even if workshop content had remained the same, concentration on the second hypothetical case study may well have proven more useful in newly-formed groups.

The "Daily Evaluation" ratings further suggest that Group D members were generally more pleased with the experience than were members of Group A. Group D either held or shared the highest mean ratings on all "Daily Evaluation" items. By the same token, Group A received the lowest mean ratings on seven of the nine items. Since workshop content was much the same for all groups, these results suggest that other variables may account for the differences. They may also indicate important variables for further study. Groups A and D were, however, remarkably similar in composition. On the characteristics of gender, employment, number of qualifications, and amount of preparatory reading, they varied little. It was only the characteristics of age and length of social work experience that showed substantial variation. The most likely explanation for the differences would appear to be that the groups' tutors affected the groups differentially. Such an interpretation is supported by the finding that members of Group D were the only ones who did not rate Thursday the best day overall. Their

Thursday ratings were noticeably lower on Items 6, 7, 8 and 9 (i.e., "adequacy of tutors", "balance in participation pattern", "discussion on topic", and "tutors' control") than all other groups. It would appear that they did not find the change of tutors to be generally beneficial. They had consistently given their tutor higher ratings than did the other groups--especially on Wednesday--and, perhaps, felt that they had the most to lose in the dissolution of the initial working groups. Interestingly, the opposite reaction appears to have taken place amongst Group A members. They had consistently given their tutor lower ratings and when they changed groups on Thursday, their tutor-related ratings rose. One could argue that the significant difference in ratings between the two groups resulted, in large part, from the differences in tutors; or, at least, from differences in group members' perceptions of their tutors. It may be significant that Group D was led by the workshop's senior tutor.

Results from the three "Overall Evaluation" instrument applications suggest that participants' reactions to the workshop experience remained relatively stable for at least five months; although there did appear to be a trend toward decline and "flattening" over time. All mean overall results registered above the graph's median, indicating that participants felt a general positive regard for the instructional program. They gave exceptionally high ratings on the "facilities" items and appeared to be least pleased with the ways in which program elements were balanced.

General comments suggest that they would have appreciated more didactic input from the tutors.

At the end of the workshop, and one month later, participants appeared to believe that they had learned from the experience. Five months after the workshop, their ratings on learning-related items remained above the graph's median but a significant decline was evident. Correspondingly, ratings immediately after the workshop indicate that the program was seen to be quite relevant to participants' on-the-job needs; five months later the ratings had declined significantly. A similar decline was apparent on the item which asked whether subjects would recommend that their colleagues attend future similar programs. Numerous hypotheses could be presented to explain this general decline: perhaps subjects simply recalled workshop details less clearly after five months, perhaps five months of on-the-job experience suggested that some of the originally-anticipated workshop benefits would not, in fact, accrue. Unfortunately, the generation of evidence to support or reject such hypotheses is beyond the scope of this investigation.

Methodology

The methods used in this Study will be a major focus of concern in Chapter VI. There are, however, a few metholodogical points which appear to warrant initial attention here. With regard to the performance test, the residential nature and short duration of the workshop suggested that a simple pretest - posttest design would represent a feasible and relatively inexpensive approach to evaluation.

The pretest segment was used to generate baseline data against which changes could be assessed; the posttest provided information on the supposedly changed state. Further, the short, residential nature of the instructional sequence made it unlikely that influences other than those associated with workshop participation would operate to induce significant and relevant changes in performance test data, thus reducing the need for a control sample. By the use of expert judges in a "blind" situation, a degree of objectivity was introduced while, at the same time, the very high expenses and intractable difficulties often associated with "objective testing" and vaguely stated goals were substantially avoided. Finally, supplementation of the performance test data with student rating information added a significant dimension to the findings. However, a number of methodological limitations and practical difficulties remain.

Performance test results suggest either that the instructional sequence did not produce significant relevant changes in participant behavior, or that the instruments used failed to register significant changes which did take place. Results of this sort occur very frequently in evaluation research on instructional programs. In this particular instance, a degree of suspicion about the instruments is aroused by certain circumstantial factors. In the first place, the tutors' instructional aims were stated in rather abstract terms. In the second place, the study was undertaken under severe temporal and financial restrictions. These circumstances combined to prohibit the development of a set of operationally defined objectives that would

comprehensively represent the stated set of goals. Thus, the construction and validation of precise and objective measuring instruments was practically impossible. In order to comprehensively represent the vague set of goals stated, a very large number of behavioral indicators--as well as a correspondingly large number of test items--would have been necessary. The resources for such an undertaking simply were not available. In fact, seldom is such a high level of resource expenditure available for the evaluation of one week-long instructional sequence. In most cases the level of anticipated beneficial effects simply would not warrant such high expenditures on instrument development. One is, then, left with the choice of either using a few limited indicators or developing some less expensive means of taking overall assessments. This Study adopted the second alternative because it appeared to offer the greater amount of return in useful information for the same level of resource expenditure. However, the low level of interjudge reliability which eventuated was not anticipated and this lack of agreement between judges tends to substantially weaken our confidence in the veracity of their judgements.

Be that as it may, none of the judges' ratings suggest that a substantial change in the expected direction did occur. The low level of interjudge reliability is, therefore, only of secondary concern. Of predominant importance is the lack of any indication that substantial positive change in performance followed from the workshop experience. Thus, the most credible conclusion would appear to be that no

significant overall change in the expected direction actually took place. An understandable lack of confidence in the veracity of the tutors' judgements casts suspicion on the quality of the information obtained, but the absence of any evidence of positive change in performance strongly suggests that such changes simply did not occur.

With regard to the two sets of rating scales, the relatively vague statement of instructional objectives led to the inclusion of a rather broad set of items. The scales' comprehensiveness would seem to enhance their apparent validity. They do not, however, represent indicators which are very direct measures of change resulting from workshop experience. That is, they measure only participant belief and opinion as expressed in one particular way; they are neither direct measures of behavioral indicators, nor apparently unbiased expert judgements. Again, the extreme lack of resources and time meant that rating instruments could not be subjected to validation studies prior to administration. Thus, the quality of the information produced must again be suspect. At the same time, the high levels of retest reliability tend to support the conclusion that participants generally perceived the workshop to be a valuable experience, even though that favourable inclination did diminish considerably with time. The important question concerns the extent to which one places trust in participant perceptions and judgements.

The utilization of a control sample in conjunction with the performance test segment of the Study was considered

and rejected for a variety of reasons; these will be discussed further in Chapter VI. It should, however, be noted that the major findings of the performance test do not suggest that inclusion of a control group would have significantly enhanced the value of the investigation. The principal necessity in a study such as this is in establishing whether or not anticipated changes took place. If such changes are indicated by the data, the control group design may provide evidence with regard to the influence of the instructional treatment in generating that change. When no changes in the level of performance are indicated, the control group design is likely to offer very little valuable information. In order to use the limited available resources most efficiently, the investigation was concentrated on gathering evidence which bore on the most immediately important issue: did change occur? It is highly unlikely that the inclusion of a control sample would have substantially affected our major conclusions. It certainly would have entailed considerable additional expense.

Finally, with regard to both the student rating data and the performance test data, statistical significance was particularly difficult to achieve because the numbers of subjects involved were very small. Had the program included two or three hundred participants, our sample sizes would probably have been considerably larger and statistically significant results more easily obtained. With only a handful of subjects, statistical significance requires extremely high absolute values.

Summary

This Study has attempted to generate information of possible evaluative use on an NISW residential workshop entitled "A Unitary Approach to Social Work Practice". It suggests that certain programatic changes may be in order and that the role of the workshop's tutors in leading small group discussions should be more thoroughly analyzed. Although various methodological difficulties and limitations tend to impair the quality of data produced, results broadly suggest that program participants were pleased with the experience. They reported a belief that the workshop had aided them in learning more about unitary approaches and their applications to social work practice. There is, however, no direct evidence to support the contention that such learning actually did occur. Performance test results indicate that the workshop brought about no significant overall changes on the dimensions studied. This contradiction in results and the unavoidable inability to thoroughly validate instruments prior to application allow us to draw only highly tentative conclusions with regard to instructional effectiveness.

CHAPTER V

STUDY II

Introduction

It is my intention in this Chapter to convey information of probable use in the evaluation of a workshop held by the National Institute for Social Work (NISW), January 5-9, 1976, at Kingswood Hall, near Egham, Surrey. The workshop was entitled "A Unitary Approach to Social Work Practice" and was led by four tutors from the NISW.[1] The information herein was drawn from a wide variety of sources, using a diversity of media which included: my own direct observation of the workshop sessions in progress, and of tutors' meetings; many informal discussions with workshop staff and participants; questionnaires administered to the participants; audio-taping of interviews with participants, workshop sessions and informal discussions; and documents provided by the NISW. It is hoped that by using a variety of information-gathering methods to tap a number of information sources, a multi-dimensional portrait of the workshop may be created. I began with close, direct observation of the workshop sessions, intending to identify and explore major areas of concern, significant issues, and perceived strengths in the program as it was put into operation. In using an investigative approach of this sort, we must depend heavily upon information which is essentially qualitative in nature.[2] Thus, it should be stated from the outset that responsibility for the collection, interpretation and

presentation of information--except for that contained in a few documents provided by the NISW--is entirely my own.

Throughout the instructional sequence, especially during the formal interviews and on the questionnaire forms, I assured respondents that their comments and observations would be presented in a manner which would attempt to ensure their personal anonymity. It is, then, not morally tenable for me to explicitly identify individuals and link them with comments or actions arising during the workshop. For that reason, no names have been used. However, in a number of instances, it is unavoidable that workshop members or tutors reading this report will be able to make informed guesses about connections between particular individuals and behavior described herein. Since identification of those connections is not integral to an understanding of this report, and since even informed guesses may be incorrect, the reader is discouraged from dwelling upon such inconsequential conjectures.

Perhaps it should be noted that this report is intended to be primarily for the use of individuals who were involved in the organization and presentation of the workshop program, principally the workshop's tutors. It is understood that the NISW intends to present similar workshops in the future and it is hoped that the observations in this report might be of some use in assisting them to reach decisions about future activities. For this reason, detailed accounts of most workshop sessions have been included. These tend to substantially lengthen the report and may be of rather limited value to those with only a general interest

in workshop processes. However, it is believed that these accounts may prove to be of value to the tutors, and that they serve as a background against which emergent concerns and issues may be explored. To the wider audience, perhaps comprised of individuals and groups employed by social work agencies and educational institutions who are presently operating or planning instructional experiences similar to the "Unitary Approach Workshop", the detailed descriptions themselves may be of little value. Hopefully, they will at least convey an impression of how the program actually operated.

Membership

The workshop membership was composed of 27 individuals, 14 males and 13 females. Their ages ranged from 25 years to 64 years and the arithmetic mean age was 38.6 years. There was no sigificant difference between the mean age of the males (38.8) and that of the females (38.3). On the whole, they appeared to be a relatively mature group of social work teachers and agency personnel; a conclusion which is supported by the observation that their reported mean length of service in the social work profession was 9.5 years. On average, males reported longer experience in social work ($\overline{X}$ = 11.3 years) than did the females ($\overline{X}$ = 8.2 years).

During the course of the workshop there was a persistent assumption, made by both tutors and members, that the membership list evidenced a disproportionately high

representation of social work teachers over social work practitioners. Closer analysis of the members' reported employment circumstances suggests, however, a more complex composite than that. Based upon their job titles and self-reported percentages of working time spent in various activities, it would appear that there were: nine social work instructors (mean age = 37.7 years), seven social work "seniors" or "supervisors" (mean age = 35.4 years), four social work teacher/administrators (mean age = 40.5 years), two fieldwork instructors, two "direct" practitioners, one agency senior administrator, one supervisor/training officer, and one full-time training officer.[3] Therefore, while it is true that over sixty percent of participants were employed as social work educators of one sort or another, it would be stretching the definition to suggest that all others were practitioners. Although these others were employed by social welfare agencies, seventy percent of them spent the bulk of their working time on supervisory and administrative tasks. They had little direct contact with persons who presented themselves as social work clients. Only the two "direct" practitioners reported spending at least one half of their working time in casework, group work, or community work activities. It is also interesting to note that of the seven "seniors" or "supervisors", only one was a female; perhaps an indication of differential rates of vertical mobility in the profession. Furthermore, it should be noted that four of the workshop participants were employed in the Probation Service; these included one "supervisor", one

"direct" practitioner, the supervisor/training officer, and one fieldwork instructor.

Members were asked to report on their formal educational qualifications and the extent of their preparatory reading. All but one held at least one formal relevant qualification granted by a recognized institution of higher education. Including "Home Office Certificates", all participants held at least one relevant social work qualification; the group average was 2.3 qualifications per person. Of the four textbooks suggested by workshop tutors prior to instructional onset, no member reported having read all of the books in their entirety. One respondent reported having read three and one half of the four, and five individuals reported having read at least two of them. Six individuals stated that they had not read any of the books at all, and the mean amount of reading reported was 1.1 books per person. The Pincus and Minahan volume appeared to have been most popular; twenty individuals reported that they had read at least part of it and eight of those twenty stated that they had read all of it.[4] Individual comments during the workshop indicated that almost all participants had read, and valued, the handout information which was provided by the NISW prior to workshop attendance.

Reasons for Attendance and Expected Benefits

Workshop participants were asked to report briefly on why they had chosen to attend the program and what, if any, expected benefits it might have for their job performance.

The responses to both questions were relatively similar, and were eventually broken down into the response areas listed below. Beside the description of each response area is a number which indicates the number of individuals who made at least one reference to that particular response area.

FIGURE 5-1

STATED REASONS FOR ATTENDANCE AND EXPECTED BENEFITS

Response Area	Number of Individuals
To gain new knowledge and a wider perspective	17
To carry new information and a wider perspective back to the teaching situation	13
To carry new information and a wider perspective back to the agency situation	8
To explore possible applications for a unitary approach	8
To learn new skills arising from a unitary approach	6
To promote greater coherance, ordering, or logic in my perspective on social work	4
To meet colleagues	4
To demonstrate my own support for unitary approaches	3
Unsure	2
To further prepare me for assisting clients	1
To help me understand others who attempt to use a unitary approach	1
To gain further ability as an instructor	1

Obviously many participants provided more than one response to the questions. It appears that the strongest "reason for attendance" or "expected benefit" had to do with gaining knowledge or "perspective", often with the intention of feeding these acquisitions back to the respondent's work situation. More than one half of the responses revolved around this concern. Interestingly, only one respondent made an explicit connection between his own learning and some direct benefit to social work clients.

Prior to instructional onset, the NISW tutors stated their major intentions for the workshop. These may be characterized as: (1) to help workshop members in exploring and understanding the basic concepts which make up a unitary perspective on social work practice, (2) to assist participants in learning to view the practice tasks associated with "assessment" from a unitary perspective, and (3) to encourage workshop members in learning to apply a unitary approach when "intervening" in problematic practice situations. When we compare members' "reasons for attendance and expected benefits" with tutors' workshop intentions, a close correspondence is indicated. Of the sixty-eight member responses, only eight do not appear to have direct relevance to at least one of the tutors' stated intentions. Although members did not respond in terms of "assessment" or "intervention" skills, their emphasis was not upon knowledge acquisition and understanding only, but also upon the application of knowledge. At least eighty percent of the responses appear to express intentions similar to those

expressed by the tutors. We might, therefore, safely conclude that the two sets of expectations/intentions were not essentially disparate.

Subgroup Composition

Most of the early workshop discussion took place in four task-oriented small groups which were composed by the tutors prior to the workshop. Figure 5-2 illustrates the resultant dimensions of each group.

FIGURE 5-2

THE COMPOSITION OF SUBGROUPS

Characteristics	Group A	Group B	Group C	Group D
Males	3	4	3	4
Females	3	3	4	3
Mean Age (years)	30.0*	45.0	38.4	40.3
Mean No. of Qualifications	2.33	2.14	2.70	2.30
Mean length of social work experience (years)	7.0	11.4	10.1	10.9
Mean No. of texts read (of 4)	1.37	0.66	1.10	1.18
Type of employment	2 seniors 2 teachers 1 tea/adm. 1 tr. eff.	3 seniors 3 teachers 1 tea/adm.	2 f/w inst 2 teachers 2 dir.pracl 1 tea/adm.	2 seniors 2 teachers ag. adm. 1 sup/t.e. 1 tea/adm.

*Two female members did not report their ages.

It would appear that the groups were composed with primary regard for the characteristic of gender. If a

relatively even mix was desired, it was achieved. With regard to the "employment" characteristic, a better mix might have been possible; Group C did not include any seniors. However, there were at least two teachers and one teacher/administrator in each group. Each group also contained at least two agency-based individuals. In light of the continuation - dissolution decisions which group members had to make later in the workshop, it might be of interest to note that Group B, the only group which later elected to remain intact, stands out as different from the others on two variables. The average age in that group was somewhat higher than the overall mean, and the amount read prior to attendance was approximately one half of that read by members of other groups. Only two Group B members had read as much as the hypothetical "average participant". Whether either of these factors have any bearing on Group B members' subsequent decision to remain together is, of course, unknown. The observations do, however, make for some interesting speculation.

Setting and Arrival

The workshop's agenda began with the arrival of participants at noon on Monday, January 5. By lunchtime, at 1 p.m., all but a few members had arrived. They had been cordially greeted by the NISW course organizer and ensconced in their allotted bed-study rooms. For the majority, who came from or through London by train, the trip was smooth and simple. Kingswood Hall proved to be comfortably "out

of the way"; allowing a degree of isolation from the outside world, yet close enough to Egham that transportation arrangements were not difficult. The course organizer appeared to be well prepared for their arrival. She obviously had established a good working relationship with the Kingswood Hall staff-person who controlled room allotments, thus facilitating simplicity in that process. She assisted individuals in sorting out their moving-in tasks and invited each one to join the others in the lounge before lunch. Furthermore, she was prepared with name tags for each participant. The value of these provisions should not be minimized. They ease the establishment of a comfortable learning atmosphere. When one is suddenly thrust into a relatively intimate residential situation with approximately thirty strangers, the warm welcome is appreciated, uncomplicated arrangements for accommodation ease anxiety, and the name tags facilitate identification. These, in turn, promote personal comfort and decreased anxiety in the initial "getting to know you" phase which is inevitably somewhat formidable.

The staff of Kingswood Hall were unobstrusive and helpful. They distributed "Notes for Visitors" which provided useful information on such matters as bathrooms, laundries, telephones, the mail, locking-up times, the grounds, car parking and room keys. Accompanying this information sheet came a map of the facility which indicated fire assembly points. This distribution of practical information in written form seemed a particularly helpful gesture in

that it not only provided easy and enduring access to information of concern to workshop participants, but also indicated a willingness on the part of Kingswood Hall staff to provide assistance when it was required.

The bed-study rooms, though not large, seemed entirely adequate. They were clean and modern; with a desk, bed, study chair, easy-chair, and adequate storage space in each. All was tidy and, although somewhat institutional in atmosphere, relatively comfortable. In all, it struck one as a decent place to spend a week.

The informal chat amongst participants before and during lunch centred around making introductions and sharing expectations of the workshop. There appeared to be an air of general agreement when one member stated: "I think I will get a lot from this course." However, even this early in the workshop, another member expressed disappointment at the small number of practitioners present. "Most of us," he said, "are either training officers or teachers."

Description of Sessions

Welcome and Introduction

The workshop's first session began on time, with the course organizer formally greeting the workshop members and providing information on such matters of interest as the availability of newspapers, evening tea facilities, coffee breaks and seating in the dining hall. Mention was made of the fact that the class had been divided into four smaller discussion groups, and that much of the workshop time would

be spent in those groups. The methods or criteria by which subgroups had been constituted were not mentioned.

At this point, the senior tutor (Tutor A) assumed control of the session. She introduced the other staff members, thanked the course organizer, and noted the tutors' intention that the workshop should become a learning milieu within which emphasis would be placed on involvement and participation. It was suggested that all members and staff address each other by given names, so as to encourage a more informal and relaxed atmosphere.

Tutor A then undertook a short explanation of the workshop program itself. This was to be the NISW's fourth workshop on "unitary approaches." The NISW was quite keenly involved in community work and had, in the previous few years, increasingly moved toward a "systems perspective." Unitary approaches had gained a great deal of theoretical impetus approximately two years previously, when the Pincus and Minahan volume, and the Goldstein text, had been published. The tutors wished to underline their expectation that the workshop would provide an opportunity for teachers and practitioners to explore unitary approaches together, and for them to share thoughts about practical applications. She then briefly discussed details of the written program, a copy of which had been provided each participant prior to arrival, and asked whether there were any questions arising from it. When none were presented, she assigned the sub-groups their task for the day--the analysis of a case study. She noted that there would be a plenary session on Tuesday

morning and enunciated an expectation that the workshop as a whole would use that time to discuss and share the small group experiences of Monday afternoon and evening. It was suggested that subgroups allow themselves time at the end of the day to prepare "group sheets" which could be hung on the walls of the plenary meeting room and used to illustrate ideas that had been discussed in their meetings. She asked, however, that verbal reports in the plenary sessions be kept relatively short, in order to leave ample discussion time. This entire introduction to the workshop was warm, but to the point. It conveyed an air of concerned professional competence.

Throughout the workshop it was, of course, possible for me to directly observe only one working group at a time. It became necessary, therefore, to concentrate my reporting on the one observed group, and hope that this would provide the major basis for an overall impression of the day's activities. It was further intended, however, that a description of the plenary sessions and the use of informants would assist in generating a more complete impression of the activities in other small groups. In any case, I have attempted to provide relatively detailed reports on what the groups did, rather than concentrating on the content pursued. It is hoped that the inclusion of these rather detailed accounts will provide a useful summarized record for the workshop tutors; a record which might be referred to when questions about workshop materials, practices, and myths are asked. At times these records include mention of

myself in the first person because I wish to emphasize that these are my observations and that I, also, was present; perhaps inadvertently influencing interaction within the group.

Monday Afternoon and Evening

During the first afternoon and evening, I observed Group C. The group's tutor (Tutor D) began the session at 2:45 p.m. by bringing participants around the table and suggesting that they spend the first ten minutes working in dyads, trying to discover each other's strongest reasons for attending the workshop. Each participant was then expected to spend one minute reporting to the group on his partner's stated reasons. The group members appeared to move into this exercise with relative comfort. Each member's stated reasons for attendance are itemized below.

Report on Member A - suggested that the "casework approach to service delivery was time-consuming and wasteful", and that he was, therefore, interested in the possibilities offered by a unitary approach. At the same time, he was somewhat pessimistic about the amount of practical value offered by such an approach.

Report on Member B - stated that he had read both Pincus and Minahan's and Goldstein's books, and had been intellectually attracted by their approach. He felt that most social service agencies and traditionally-oriented schools of social work were in need of

some change in outlook, toward a more "integrated" perspective.

Report on Member C - said she had just completed a two-day course on the implications of a unitary approach to fieldwork placement and was very interested in further exploration.

Report on Member D - in her role as a teacher of social work, she was moving toward instruction from a more unitary perspective. She was, however, rather skeptical of the "bandwagon" element in adopting a unitary approach.

Report on Member E - stated that the training officer in her agency had suggested that she attend because it might help her to develop "better supervisory abilities." She knew very little about unitary approaches.

Report on Member F - this report was terribly muddled and needed constant correction. Her reasons for attendance were not clearly presented but it was clear that she had read some of the literature on unitary approaches and had not yet tried to put any of the theory into practice. Later, it became clear that she worked in an agency where implementation of a unitary approach was being seriously considered.

Report on Member G - felt that he, along with many of his colleagues, had developed an "intellectual acceptance"

of unitary approaches, but had not yet developed any "feeling acceptance" of them. He hoped that the workshop would help him to integrate theoretical underpinings more deeply into his personal value orientation.

This exercise appeared to be a useful icebreaker. At the same time, it opened the door on a few of the issues which would continue to be of concern throughout the workshop. It was, therefore, judged to be a good beginning.

Once the exercise was completed, Tutor D shared with the group his notion of what a workshop should be. He emphasized the adjectives "democratic", "participatory", and "collegiate". He outlined his perception of the group's resources: basically, the members' attributes and the materials which had been prepared for them. He then asked that they establish a contract with him, noting that he would attempt to help the group in collectively agreeing upon tasks and that he would try to keep the discussions focused on the subject matter that was agreed upon. He did not wish to be seen as an "expert" on unitary approaches but said that he would contribute what information he could, act as a timekeeper, and serve as the group's link with the "management" of the workshop. The group members asked a few questions, seemingly to clarify the "contract", but did not explicitly accept or reject it.

Tutor D introduced the "Whittingworth Buildings" case material and reminded the group of the expectation that they would make some presentation to the next plenary

session. He suggested that it might be a good idea to select a group recorder, and perhaps a spokesman. However, this suggestion was not taken up at the time. There was some questioning and discussion with regard to the origins of the case material. It appeared that most group members were somewhat unsure or unwilling about starting an analysis. Tutor D suggested that they establish a way of working on the material, and one member suggested that the selection of a group recorder might be a worthwhile point of departure. Another argued that no decisions should be made about the form of their presentation until the group had become more involved with the material. The tutor then suggested that each member keep his own record of the group's activity and that later in the day they could synthesize from these. Members agreed upon this as a temporary course of action. The tutor then asked, "What are the major difficulties in this case which face the area team?" The group responded with a tentative exploration of the material, talking on a highly intellectualized and somewhat superficial plane. As yet, they had produced no conceptualization of a format which might systematically guide their work. They continued to operate at an apparently superficial level for some time, until a few members expressed frustration at their obvious lack of progress. It was noted that the case presented several problems which required interventions by a variety of service agencies. This observation was not, however, taken-up and one had the feeling that group members were not really willing to become intensely involved in the work.

I found myself becoming bored and drowsy.

Finally, Tutor D challenged, "You are now the area team charged with handling this case. What are you going to do with it?" It occurred to me then that the group might have started to work much more readily had the exercise been more clearly presented as a simulation, i.e., had the tutor's question been asked in that form, at the beginning of the session. Tutor D reported upon his perception that members had been relating to the material on a very superficial level. He suggested that they "forget the past, begin again." One member expressed frustration at the group's apparent unwillingness to become involved in the case situation. The tutor, again, said, "You are the team; how do you begin?" One could feel frustration mounting in the group as the tutor prodded once more, "You are scared stiff of starting." He acknowledged that there were differences amongst group members, in the knowledge base that each brought to the workshop, but exhorted, "Let's not allow that to cow us into silence."

At this point, Member B volunteered to act as recorder while the group made a "laundry list" of the problems noted in the case material. This appeared to give them a concrete starting point and the expression of ideas began to flow more freely.

It took over an hour for this group to really begin work on the exercise. Tutor D prodded them consistently, but it was not until a group member took the initiative to provide some leadership that any significant task-oriented

discussion took place. Once they were moving, Member B assumed an even more directive role; not only by using his power as recorder, but also by offering alternative plans for the analysis. He was joined by Member A who had been troubled by what he said was a need to make "arbitrary assumptions" about the case situation. Although his contributions to the discussion were somewhat obtuse at times, he did stimulate other group members to act. At first, his presence appeared to intimidate the rest of the members. His comments were critical and his manner always argumentative. But, as they became accustomed to him, the others began to express their own opinions.

The tutor pointed out that this case seemed to "overwhelm" the group, and Member A made another practical suggestion with regard to a means of approaching the material. By 3:55 p.m. the participation pattern, though far from intense, was relatively homogeneous--except for Member G who had said almost nothing. The tutor had retreated from his acutely energizing stance. Member B continued to assume a leadership role and by teatime, just after four o'clock, some of the frustration had obviously eased. Before adjourning for tea, the tutor attempted to set a distinct task for the group to begin when it returned to work. He suggested that the session, thus far, had not provided them with a clear idea of means by which they might best approach their tasks and noted that the question "Where do we start?" continued to require an answer.

During the tea break, by informally chatting with various members, I attempted to gauge the states of affairs in other small groups. Almost all participants were entirely willing to provide an analysis of what had occurred. It appeared as if Group C was not the only one to experience difficulties in coming to grips with the case material. There was, however, a general consensus which indicated that an impetus toward task achievement was beginning to develop in each group. All were reported to have begun with expressions of subjective impressions, "pulled from the air", rather than any search for procedural directives. Tutor B noted that in her group (Group A) one of the members had started the discussion by outlining exactly what he would do about the case situation. This gave other members a stimulus to which they could react. Intense involvement soon developed. Hers appeared to be the most vigorous group at that time.

At 4:35 p.m. Group C resumed work. Member A immediately began to demonstrate his facility with unitary approach terminology, prompting Member B to ensure that all other group members possessed some basic understanding of the concepts under discussion. His direct question occasioned a somewhat unconvincing nod of heads. The general discussion resumed, but it was almost entirely confined to interaction between the tutor, Member A and Member B. Member A, in fact, had become very aggressive in making his points and, although most of his comments were rather ambiguous or illogical, no one expressed any lack of

understanding. The tutor found that he had to prod the group continually and, by 5:05, any impetus that had previously developed was lost. No one ventured any comment on the processes occuring within the group. No one seemed willing to say, "What are we doing and what should we be doing."

Finally, Tutor D asked if any of them was willing to outline his own perception of one major problem in the case and propose a solution. Member B again took the initiative. Member A, however, became extremely obstructionistic. He did not think the case analysis could proceed without further detailing of information. Member G followed B's lead by suggesting a specific action that might be taken by an agency team which was attempting to deal with the Wittingworth case. Member A became even more hostile and was met with growing expressions of resentment from other group members. Expressed in terms of "Games Theory", Member A played a very hard game of "Yes - But". He continually used his perception of the "reality of the situation" to stymie any impetus toward a firm commitment on starting at a particular place. This appeared to prevent the group from actually dealing with elements in the case situation. They were unwilling to talk about specific action plans, ostensibly because the case material was presented in "over-generalized" terms.

Member F then commented on how she saw their corporate behavior. She reported her own frustration at the fact that whenever anyone made a procedural suggestion, he

was met with silence, sleepiness and yawning. Her comment was met with silence, sleepiness and yawning.

By this time it was 5:15 and the tutor tried again to prod members into action. He suggested that they take more risks by making more definite statements. Perhaps, he said, they could begin to more thoroughly explore a few alternative interventive modalities and deal with the attendant risks later.

The multitude of problems presented in the case study appeared to combine with Member A's hindrance to overwhelm the group. Members had reached the point of making absolutely no new suggestions, probably because they knew that Member A would always have "reality factors" to contradicate their suggestions. His cynicism appeared to overpower all attempts at task accomplishment. It became certain that if he continued, sooner or later someone would have to find a way of dealing with him directly. No one, however, had verbalized this feeling as yet. The group was in trouble. One could feel their frustration, produced by the knowledge that they had accomplished very little and that time was running short. To a large extent, Member A controlled interaction in the group. I noted, for example, that the three members seated directly across the table from him contributed almost nothing. One could not help but wonder if his eye contact might have been a reason for their silence.

At 5:35 Tutor D again pushed them to arrive at a procedural decision. Member B took up A's proposition

that they were having trouble in decision-making because there was a lack of information in the case material. He said, "We get to the point of 'either-or', but cannot decide on an alternative because we lack information." It was then suggested that they look at some of the more long-term changes which might be initiated in the Whitting-worth Buildings environment. A disjointed discussion ensued.

After approximately ten minutes, Member B attempted to draw Member G into the discussion by asking for his reactions, as the only probation officer in the group. G responded and the discussion intensified quickly. B then hit upon an imaginative idea which concerned one possible interventive tactic. Basically, the idea rested upon cooperation between the area social work team and the local probation office. The tutor encouraged further development of the suggestion and Member B presented some possible refinements. The discussion had suddenly become much more lively but, unfortunately, time had run out. The session continued until after six o'clock and Tutor D brought it to a close by setting a specific task for the after-dinner session.

The group reconvened at eight o'clock with Tutor D reiterating the tasks which he visualized for them in the next hour and a quarter. Members appeared to feel more attuned to one another--more "chatty". Member A, however, was not present. The tutor asked, "Does anyone have any idea of how to tackle these tasks?" and Member B made a procedural suggestion. The tutor modified the suggested

approach slightly, introducing a recommendation for the division of labour. At this point, Member A arrived, prompting B to summarize the discussion thus far. This led Member C to suggest further modifications in the plan of attack. B continued to maintain a directive stance and almost forced the rest of the group to commence work.

At this point the tutor decided that he would leave group members to work on their own. He did not explain his reasons for so doing. After he had left the room, one member made a comment about his having broken the contract. This elicited a few giggles. With the tutor away, more noticeable attention was paid to my presence; but the group was working. They tried to set down on paper the major problems in the case situation, and this became a relatively light-hearted task. Member A, however, continued to cling to a censoring role. At times, he appeared to hold vetoing power.

Member G became much more involved in the group's work. He reintroduced the possible influence of the probation office in the case and attempted to develop specific interventive strategies which included action emanating from the Probation Service. He noted that many probation officers did not enjoy being seen as different from social workers. "Probation," he said, "is more hidebound by tradition, whereas social work is by administration."

As work continued, the group began to apply increasing amounts of theoretical material from the literature directly to the case. They referred to Goldstein's diagram

of a unitary approach and wondered if the case circumstances might be presented in similar manner. Member C took charge of the pens and paper, and began trying to detail it out with them. Some members were, however, obviously tired. There was much yawning and a number of participants tried to revive themselves by walking and stretching. Member A continued to play "Yes - But", but in a lighter vein.

At 8:45, Tutor D returned to the group. Someone made the comment that they were all rather weary, but that their presentation to the plenary session would be awfully dull if they did not make substantial progress in the twenty minutes which remained. The tutor tried to revitalize them by suggesting that they break into dyads to discuss various segments of the case. Member D expressed annoyance at the fact that they had not done as much as she would have liked and the tutor pinpointed the need for a "clarifying mechanism". Both comments, however, were met with apathy. They thought that the multitudinous factors in the case had overawed them, and that the case material was somewhat "unreal".

Member B suggested a means by which they might present their work to the plenary group without having to do any elaborate preparation. His proposal involved the use of a "fishbowl technique" whereby they would hold a pseudo-discussion in the plenary session and bring about a "spontaneous output of feelings about issues". The tutor suggested that they might at least consider drawing up an agenda of issues to be covered. For the last ten minutes,

two discussion topics competed for favour. One concerned means by which they would make their presentation; the other involved seeing "change agent systems" as "targets".

By this time it was 9:10 p.m. and most of the participants seemed to be only half awake. There was, however, still nothing on the blackboard to indicate a summarization or a plan for the impending presentation. The tutor asked them to make a decision about how they were going to present their material. They decided to use the "fishbowl" and spent the last ten minutes drawing up an agenda of issues for presentation. At 9:25 p.m. the session ended.

Apparently, the three other subgroups also spent a good deal of time establishing themselves as integrated entities and discussing tasks which applied to analysis of the case material at hand. By the time the first day's meetings were completed (9:15 to 10 p.m.), many group members reported weariness. Between travelling, moving in, personal introductions, the introductory session, and approximately four and one half hours of small group work, they had completed an exhausting first day.

Monday Tutors' Meeting

The day's work, however, was not over for the tutors. At 9:35 p.m. they met to discuss the events of the day and to finalize plans for Tuesday. Tutor C arrived a few minutes late, apparently because his group had continued to work until after 9:30. Each tutor reported on his or her group.

Group A

Tutor B remarked that members of her group did not seem to be "remarkably innovative", but did appear to be relatively competent and knowledgeable about unitary approaches. They had demonstrated a strong identification with practice principles and had been led into their work by a group member who was highly involved in implementing a unitary approach on the job. The tutor felt that her group could have profitably used another day on the Whittingworth Buildings material.

Group B

Tutor C expressed disappointment with his group. He said that they had clung to the erroneous assumption that the practitioners in their midst were "automatically" already using a unitary approach. He found many of their comments facile. Further, he was concerned that the textual material provided prior to the workshop had not been widely read. This, he said, had led to much "groping about"; attempting to grasp the conceptual models. The discussion had not been "tied up", rather it was scattered and almost incoherant at times. Furthermore, he said, the powerful male contingent overshadowed the females. In his view, the group had not considered many of the problems that might have been identified in the case study, although they did display a willingness to argue points out. They had also voiced a strong rejection of the tutors' expectation that they would report back to plenary sessions. On the whole, he felt that Group B had terminated the day with a good

deal of "unfinished business". They had ended without developing any plan for reporting to the plenary session "because they wanted to retain the 'bits and pieces' nature of their discussions."

Group C

Tutor D characterized his group as "five teachers struggling with the practical problems of a unitary approach." One member was constantly involved in doing "heavy demolition jobs" on other members' contributions and, perhaps because of this destructive element, the group members seemed diffident and reluctant to contribute. At times Tutor D had felt that he needed to almost bully them into action. He was impressed with an observation that the "difficult" group member behaved as two different people: one, vivacious and friendly during the dinner hour; the other, cynical and obstructionistic within the small group. On the whole, the tutor was not very pleased with Group C's accomplishments.

Group D

Tutor A characterized her group as "thoughtful". One member had been particularly silent; he may have been lost in the theory. Another was quite dominant and pontificating, but seemed to realize this at times and, all in all, was not disruptingly powerful. The group as a whole appeared to be "too comfortable" at times. However, they had been "galvanized" by one member who was willing to openly state his opinions and involve others in discussion.

Other Comments

Some problems were being encountered because a few names had been misspelled on the group lists. This was being corrected.

There was some discussion over the tutors' perceptions that quite a number of the participants did not seem able to "embrace" the case material very well; many "objectified" it. It was thought that a large number of the members had little or no previous experience with this sort of material.

They then examined the things which had to be done on the next day and delegated responsibility for each task.

This group of tutors impressed me as having done it all before. They seemed competent and comfortable with each other. In fact, it appeared that each drew some strength from meeting with the others. Tutor A obviously held the reins of the workshop, but was very open to discussion and democratic decision-making. She appeared to maintain control because the others respected her, listened to what she said, and allowed her opinions to influence them.

Tuesday Morning Plenary Session

The session began punctually at 9:30 a.m.. The walls of the room were littered with "wall hangings" which illustrated the tortuous discussions of the previous day. The participants, however, did not seem to have enough time to digest them completely; even by 9:45 it was apparent that no more than sixty percent of the class was paying

attention to Group A's reporter. The rest were still involved in reading the illustrations.

Group A finished its seven minute report and Tutor D, the discussion leader, threw the meeting open for discussion. After approximately one minute's silence, Group A members began to take up some of the points mentioned by their reporter. Tutor D encouraged members of other groups to respond to Group A's analysis. An inhibited discussion ensued. Eventually, it was Tutor D who outlined some of the shortcomings of the analysis. Group A members were, however, able to defend their work well.

When asked to report, a member of Group B stated that the group members had been under the impression that they would not be asked to make a presentation. He suggested that the expectation was viewed as an inhibition to learning. Another member took it upon himself to make an informal presentation, even though he stated that he tended to agree about the inhibiting effect of expecting a report. After his short report, Tutor D encouraged discussion on the subject at hand by asking for reactions from other Group B members; even though one Group B member had already suggested that they move on to hearing Group C's presentation. Fragmentation within Group B was evident and the discussion continued, half on subject-oriented issues and half on expositions of group dynamics. Group B appeared to have done a great deal of exploration but had reached very few resolutions from their deliberations. Be that as it may, the informal, spontaneous presentation was forceful and

interesting. Ensuing arguments gained attention from the whole class.

The members of Group C attempted to present their material in the form of a simulated "fishbowl discussion". As the argumentation intensified, the entire membership became involved in exploring a number of "tensions" in the case situation. The "fishbowl" turned out to be a worthwhile demonstration of some team problems that are likely to arise from adopting a unitary approach. General discussion became so involving that Group C had difficulty in bringing its segment to a close.

The entire class was attending intently when Group D began its presentation. Points were taken up and explored quite thoroughly at times. However, it became increasingly obvious that many participants were rather naive about some of the definitions and principles of "systems theory" which were bandied about. There was no consensus on a meaning for the word "system" and some participants requested more didactic input on the subject. This interest seemed most appropriate at the time because the following session, planned for after the coffee break, was to be Tutor A's talk on various conceptualizations of a unitary approach. Whether by accident or design, members appeared to be ready for such a presentation.

Tutor A's Presentation

The senior tutor began by drawing member's attention to the handouts which had been prepared for them. Her talk,

she noted, was to be concerned with exposing some of the salient features in the various unitary approaches to social work practice. The discussion would not be aimed at direct exploration of "systems theory" as such. It was her hope that an exercise which had been planned for later in the day would assist participants in gaining a deeper understanding of "systems" concepts. She then presented her own distillation of the principal ingredients of a unitary approach and guided the class through a brief exploration of differing perspectives on the subject. It was a well organized and concise exposition. The impression gained later was that workshop members had appreciated it immensely. Her use of transparancies on the overhead projector emphasized major points and assisted those who wished to take extensive notes. At one point, however, she said, "If you have the paper on Whittaker in front of you" and was rewarded by every class member conducting a search through papers and folders. This is, of course, a very small point, but it was entirely avoidable; there was no real need for them to find the Whittaker paper. The change in activity distracted members from the notions she was attempting to convey.

Before her presentation was complete, a tangential discussion developed. It was primarily concerned with various problems involved in teaching a unitary perspective. Although it appeared to be of interest to most participants, it was not especially relevant to the topic at hand. Certainly, as they moved toward a closer examination of the <u>utility</u> of a unitary approach, almost all members became more deeply

involved. The discussion roamed over a relatively broad area for approximately forty minutes and appeared to reach an apex when members explored the values and priorities of social workers, relating them to possibilities for a closer liason and clearer information flow between social service agencies and educational institutions. The group finally terminated its own discussion in order to hear the rest of Tutor A's presentation. She finished off quickly and asked members to continue their consideration of issues which might have significant influence in the workshop during the rest of the week. She offered notes on the Pincus and Minahan text and handed out the exercise which groups were expected to deal with during the afternoon session. Tutor C mentioned that the exercise had been "broadly drawn" and that there would be a briefing session before tea, so as to promote a better understanding of it. Most members, however, were still deeply involved in discussing points arising from the morning's session.

The mood of the workshop at this time appeared to be especially positive. Tutor A's session had allowed for a good deal of questioning, explanation, and ventilation of feelings. Everyone, however, appeared to be ready for the scheduled free early afternoon. It was evident that some participants were feeling tired. Through the lunch hour individuals were commonly heard to say that the workshop was extremely involving but exhausting. This lunch period and free early afternoon were generally regarded as signifying the end of the first workshop segment. The majority of

members spent their free time walking together on the land surrounding Kingswood Hall. Some of the tutors joined them.

Tuesday Afternoon Groups

The briefing session which was held before sub-groups began work on the new task, emphasized that the purpose of the exercise was, primarily, in examining answers to the question: "What do we mean by 'system'?" Basically, the idea was to use the workshop itself as a subject for analysis. The time limit on the session was extended by one half hour, in the belief that the task would require a great deal of work. Tutor C attempted to clarify expectations and received little response to his request for questions of clarification. This was to be the tutors' first experience with that particular exercise and they had decided to operate only as consultants; removing themselves from the groups to do their own analysis of the workshop, but remaining available to provide consultative services should they be requested to do so.

Since this exercise had not been used previously by these tutors and since the tutors did not attend subgroup meetings during this session, it seems particularly important that this report should contain a relatively detailed description of events in the observed group. Again, the description is taken directly from my notes which were made while the session was in progress.

I observed Group A during the whole of this exercise. The reader may remember that Tutor B reported them to be

"relatively competent and knowledgeable." Although this was not immediately obvious, because the first two procedural suggestions were not taken up, it soon became apparent that the members had developed a real ability to work together. The first two suggestions were: "Let's elect a group leader," and "Let's elect a recorder." No member volunteered to stand for either of these posts and it was decided that they would try another approach. It had been argued that a leader or recorder might not feel free to contribute fully to the discussion. They therefore decided to leave the roles formally unfilled and reserved fifteen minutes at the end of the session for transcribing the discussion points into a form suitable for presentation. They decided that by using the blackboard to note and order the points as they arose, they would not only have a simple recording of their deliberations but also a focusing mechanism. As the session went on, this procedure appeared to work very well; although at various times the group wished it had some three dimensional means by which their analysis could be recorded and exemplified.

The analysis began with identification of the various subsystems within the workshop. Interpersonal interaction was close, friendly and spontaneous. Individuals appeared to feel free to use the blackboard in order to visually illustrate their points. Discussion became very brisk. At one point, the interaction was so hectic that I wondered if Tutor C had overemphasized the time constraints. The important questions: "What do we mean by 'system', and

should we try to define it?" were left with no initial answers. Some frustration about definitions was expressed. It was pointed out that the workshop tutors had not presented any sort of answer to these questions and that, although this may have been intentional, it meant that the group had to work very hard, expending a great deal of precious time on "the preliminaries". One member kept suggesting that she get her books and that they should appeal to the authoritative literature for definitions. This member was trying to move toward task completion very quickly and was almost dragging the rest of the group with her as she rushed through various conceptualizations. Two other members, however, were conspicuously uninvolved; at least, not participating verbally. The rest of the group seemed to accept this behavior without comment.

The participative group members moved freely back and forth between the blackboard and their chairs. They did not appear to notice my presence until someone noted that "the evaluator" should be seen as a subsystem within the workshop. As they began to consider it, they were rather unsure about my function and the "environmental presses" which impinged upon my role performance. Eventually, they worked it out to their own satisfactions and moved on to pay complete attention to another subsystem. I reached the conclusion that my presence was not substantially influencing their behavior. This sequence, in fact, exemplified the procedure they followed. They would identify a subject for consideration, concentrate fully on it until

they were satisfied with their answers, then identify another subject for consideration. They demonstrated an extremely strong commitment to task completion.

The group suspended its activity for a short time while one of the most inactive members finally went to fetch the textbooks. However, one could see the impetus toward task completion growing again even before he returned. The bulk of the group was always concerned with "moving on" and by the time the books had arrived, discussion was making headway again. In the end, the books were hardly even opened.

A few minutes later, the group was interrupted by an emissary from Group C who asked if one of them would accompany him, to tell his group what they had been discussing. He said that his group was somewhat "lost" and was looking for direction. The members of Group A discussed his request for a minute and decided to refuse. Sending a delegate, they said, would "disturb their flow". There also appeared to be an undercurrent of feeling that the four small groups were engaged in a competition with each other. Group A was not prepared to help the "opposition". After Group C's representative left, there was some discussion of his possible motives; however, the members soon dropped this topic in order to return to work.

It did not take them long to again become totally immersed in the analysis. Some members were delegated to transcribe previously rendered schemata onto wall-hanging sheets. Others continued to develop new ideas. With some confusion, they managed to complete a great deal of work in

very short order. However, because they tried to include so much material, their "systematic analysis" looked as if it was going to make for a rather confused presentation. The energy level had remained high and the "rushed" atmosphere continued to swell. It was suggested that the group might consider working through the dinner hour in order to complete a clear analysis and presentation. At 6:05 they took up the suggestion and decided to postpone any decision until 6:30.

During the ensuing discussion, a point of particular evaluative interest emerged. Group members were discussing the relative utilities of the various subsystems to the overall accomplishments of the workshop. There appeared to be a strong consensus that the greatest contributions to student learning had come directly through discussions with other participants. Members felt that the tutors had, undoubtedly, provided useful contributions, but they were agreed that the workshop structure--encouraging involvement between members--had proved most valuable.

By 6:30 they had tacitly agreed to finish by dinner time. The work was going well and new ideas kept sprouting forth. More strongly than ever, there was a feeling that they were in competition with the other groups; they wanted to put on a "good show" at the plenary session. Finally, they decided to call a halt to the development of new ideas, in order to organize their presentation. By this time, leadership in the group had been shared quite equally amongst all but one member. He had contributed very little during

the entire discussion and was not put under any pressure to do so.

It was decided that the member who had been silent earlier in the discussions should be the group's spokesman at the plenary session. He had become much more active in the latter half of the session and appeared to provide a stabilizing, logical influence. The other members felt that their presentation would have more impact and consistency if he was to lead it. They therefore spent the last ten minutes summarizing their discussions and helping the spokesman to complete his notes. The one remaining silent group member took on the physical labour, so to speak. He carried each wall-hanging through to the plenary room and arranged the visual presentation. No one else left the room until the spokesman said that he was prepared enough.

My own feeling after this session was: "What a lovely group." They demonstrated a very high commitment to task achievement at all times and operated with a consistent positive regard and courtesy for each other. The session had been exhausting but, at the same time, really exhilarating.

It should be mentioned that other workshop groups were reported to have had very dissimilar experiences. One was reported to have been "very depressing". Members said that it had "regressed into numerous power struggles" and had taken a great deal of time in accomplishing anything toward task-completion. Another group was reported to have become very "encountery", i.e., its collective attention

was preoccupied with the interactions occurring within itself. The member who made that observation was obviously shaken and, perhaps, rather angry about the experience.

Tuesday Evening Plenary Session

During this evening session each group was allowed ten minutes to present and explain its own analysis of the workshop. It was intended that this would assist participants by generating deeper understanding of systems theory. Participants seemed to be most interested in seeing how other groups had perceived the workshop's structure, and how each had organized its enquiries. One comment from a Group D member is illustrative of the major problem encountered with the exercise. He said, "We want to know better what is meant when we talk of a "client system". In fact, there were numerous comments which indicated that the concept "system" continued to be ill-defined and poorly understood. Members appeared to perceive this as a very pressing problem. There were suggestions that it might have been wise for the tutors to have presented a short didactic exposition on the various conceptualizations of "system". It appeared as if the exercise, itself, had done little to promote basic understanding of the concept. Other alternative means were not discussed.

All four groups reported that they had spent the afternoon deeply involved in discussion. One group, however, suggested that they might have stayed closer to work on the assigned task had the tutor remained in the group to focus

them.

Group A presented a very complex yet coherant analysis. They put on the "good show" which they had intended, and the workshop membership as a whole paid very close attention. In fact, throughout the plenary session, attention and involvement remained high. During Group A's presentation, it was strongly suggested that the workshop needed more participants with "a community work point of view".

Group B's presentation produced a signal comment from a Probation Service supervisor. He stated, quite vehemently, that the afternoon session had been a "milestone in the course". Prior to the "systems exercise", he said, he had felt separate from other members. However, during that session he began to feel that he was contributing something of value to the group, and that other members were contributing directly to his learning. He reported, "We got close; I didn't feel outside." At that point, an uninvolved observer might have posited that by Tuesday afternoon group members had been together for sufficient time to permit the development of strongly cohesive ties--especially when members had been so heavily involved in working on a shared task. This cohesiveness, however, did not seem to have taken place in at least one of the groups. Without a tutor to "focus" them, Group C members agreed that they had become somewhat lost in the material and that there had been a breakdown in some of the previously established interpersonal relationships. Be that as it may, of Group B's

experience the Probation supervisor said, "This was the right time for us to do this particular exercise . . . it was meaningful."

Tutor C attempted to amalgamate some of the discussion points that had been put forward during the presentations. He asked, "What has been clarified by the exercise?" It was generally agreed that the session had assisted participants in seeing that, as one member put it, "things are parts of other things"; i.e., it helped them to view sets of phenomena from a "systems" perspective. It appeared, also, that members had concluded that the primary focus in social work practice should be upon the "totality of the problem", rather than on any specific ingredients which might arise from a casework, group work, or community work orientation.

Perhaps one final comment on this session is in order. It was during this plenary session that the concept "client system" became corporately recognized as a more general abstraction than "client". In its new usage, members felt that they had become more alert to components of situations that might otherwise have gone unnoticed. The "client system", therefore, was seen as a much more useful tool, even if understanding of it remained somewhat limited.

Wednesday Morning Group

On Wednesday, participants continued to work in those same small group configurations until 4 p.m. Assigned tasks centred around the Rimington case material and the

major objective was for members to gain some experience in organizing agencies and services with a unitary approach in mind. Discussion began slowly in Group D, primarily because of difficulties encountered in the case material. There appeared to be a belief that the Study was too similar to the one used on Monday. Members complained that it did not contain explicit enough information to allow for task accomplishment. Tutor A suggested that they could add material if they wished and clarified possible approaches to the work at hand. Members set about rereading the case material and someone commented that it seemed like "an awful lot". They then settled down to making decisions about how they would tackle the project and were soon involved in a general discussion of the case's most striking features. They worked together until the coffee break at 11:05. A high level of cooperation between members was evident and, collectively, they demonstrated good analytical skills. There did not, however, appear to be a great deal of vitality or enthusiasm in the group. They functioned without any real leader, although the group member who wielded the pen as they diagrammed out their ideas did exert a measure of control over the discussion. Tutor A was left free to comment at any point and she did register her perceptions with them regularly. Once an overall systematic analysis of the case was complete, she suggested that they get on with developing a general strategy for planned change. They began to work on this just prior to coffee time.

At the coffee break, members of other groups were questioned about their morning's experiences. A Group B member reported having had a hectic time but said the group had "remained basically comfortable". He thought that Group B members would probably decide to remain together, rather than disband and reconstitute into new groupings on Thursday. A Group C member also reported having had a very intense morning, while a Group A member said that her group had been "pretty dead". It was at this point that I made my first request for a formal, taped interview--with one of Group C's members who had reported having difficulties with other group members. The individual politely rebuffed my request, stating that she was "not ready to risk yet" but that she would reconsider the request in a day or two. As it happened, this was the only workshop member approached who did not consent to a formal interview.

After the coffee break, I returned to observation of Group D. Members decided to partialize their common task and designated two subgroups to deal with the resulting segments. One member was designated "liason officer"; responsible for promoting a flow of information between the two subgroups. The tutor volunteered to find more paper and pens, so that members did not have to interrupt their discussions. Then she and the liason officer moved freely between the subgroups, cross-pollinating ideas and encouraging refinements. At 12:05 the subgroups amalgamated and the former liason officer took on the role of group chairman. He outlined the subgroups' accomplishments and led members

into a discussion of "action plans". During this episode, Tutor A performed a very useful service by questioning and focusing the group's efforts; however, she seldom imposed her own formulations into their scheme. At 12:40, they adjourned for lunch.

Conversation over the luncheon table did not dwell on the content of the morning's deliberations. Two comments were overheard, both from social work teachers, to the effect that the workshop was proving to be a success because: it was a break from routine, and it provided an opportunity to gain colleague support for innovatory planning. This second area of benefit, they felt, was undoubtedly the most important part of the experience.

Wednesday Afternoon Group

Since it was the only working group which I had not yet observed directly, I spent the afternoon with Group B. This was to be my last opportunity to see them in action before the groups disbanded and reformed for the Thursday sessions. They began work at 2:40; approximately ten minutes late. Various members entered the room, sat silently, left and returned. No one appeared willing to begin until all members had assembled and were ready for work. Their deliberations moved very slowly at first. No one evidenced any strong feeling about directions in which their work should be headed and one member appeared to disrupt any continuity which began to develop. He would sit silently for a few minutes, then suddenly burst into the conversation with

points which often appeared to be entirely irrelevant. The other members, however, displayed great tolerance of his behavior and I came to gain the impression that he was valued as an innovator; the group's "artist", so to speak. For over half an hour the "discussion" continued, seemingly accomplishing nothing toward task accomplishment. From time to time, Tutor C reminded them that working time was limited. Finally, at about 3:25, the group's most steadily logical member expressed her frustration with their lack of progress. She said that she wanted to stop "just going 'round and 'round" and try to accomplish something. The group's "artist" immediately challenged, "I will learn less by trying to arrive at resolutions." The frustrated member responded, "I disagree, I think we will learn more by trying to complete it." The "artist's" rejoinder was, "I wish I could be upside-down." On the surface, the "artist" appeared to be disruptively superficial; partly because he was loud. However, at another level, he appeared to play a useful, imaginative role. Most other group members seemed to realize that. His systematizations were often quite solid and he had the ability to radically alter perspectives, introducing new, yet relevant, ideas. By 3:30 each group member had expressed his or her own perspective on the task, and all were contributing freely.

It was during the ensuing discussion that I again witnessed a valuable characteristic of the participative workshop experience. On a number of occasions during the discussion, one member would tell the rest of the group about

something that had occurred in his work. Usually, these anecdotes were accounts of some event that had been viewed as a valuable change in procedure. Often they were conveyed in the form: "This is what we did and the consequences were . . .." Such interjections were usually relevant to the topic under discussion. One was impressed with the likelihood that such colourful anecdotes would be stored away in participants' memories until some appropriate on-the-job conditions elicited them. It seems very likely that vividly described, practical suggestions--arising from colleague's experiences--will be carried back to performance on the job. This would appear to be a specific example of useful "cross-fertilization" between workshop participants.

By this time, there had been a great deal of verbal activity in the group; much of which appeared to make no contribution to task accomplishment. Yet, it was not until only minutes before the designated quitting time that the level of frustration became high enough to force consideration of the impending plenary session. One member asked Tutor C if he thought the group had come near to accomplishing its task. The tutor replied that the previous ten or fifteen minutes' work had contributed a great deal to task accomplishment but that it had also cut off valuable discussion on another topic.

The group's "artist" stated, "I am left with a great sense of vagueness." He proposed that, rather than disband, the group should remain intact and apply its "newly-found systematic approach" to the case material on Thursday. He

believed that a "false focus on task accomplishment" had prevented them from fully exploring their ideas; further, that the group was now much more able to clearly focus than they had been on the previous day. Some group members hurriedly agreed with the proposal; a few did not appear to be so confident. Tutor C interrupted, asking what they intended to do at the plenary session. A discussion "starter" was quickly designated and members said that they would "play it by ear from there".

The "artist" then returned to his proposal. "Are there any dissenting voices in the group," he asked. No one responded and it was assumed that they had agreed to remain together; all, that is, except Tutor C. The "artist" said that he would "make a bid" for having the tutor join them again. He felt that group members had gained a very great deal from the tutor's presence. The session ended fifteen minutes into coffee time.

The plenary session began at 4:40 p.m., after allowing ten minutes for viewing of the wall-hangings. Three of the tutors arrived slightly late and appeared to have just completed a hasty meeting. Tutor B, the discussion leader for this session, began by setting the same ground-rules as had been used in the previous plenary sessions, i.e., ten minutes per group presentation. She asked, however, that the last fifteen minutes of the session be reserved for constitution of the new small groups.

The entire class appeared to be very attentive to group reporters. Again, with each group having had the same remit, they were most interested in determining how other

groups had approached their work; and in the accomplishments which were reported. During Group C's presentation, which was structured as a role play, it became obvious that one member held strong negative feelings about the usefulness of participating in that group. This was the member who had earlier rejected my request for a formal interview. Again, I thought that her comments might be valuable. The role play did not appear to satisfy Group C's intentions for it and one participant from another group--the "artist" from Group B--commented that Group C did not have anything new or useful to offer. He allowed that he might, somehow, have missed their point, but the class as a whole appeared to be in sympathy with his comment. No one was very impressed with the presentation. However, I also got the impression that Group B members were actively trying to devalue the work done in other groups; perhaps in an effort to minimize their own lack of task accomplishment.

Group A, again, was most impressive; moving directly to the introduction of important points and dealing with them in fine style. I could not help but believe that this had been the best of the groups. These individuals had worked hard on the assigned topics and, yet, had reserved time for thorough discussions of their own group structure, and dynamics. As a result, they had produced some excellent observations. It had become clear to them that commitment to a unitary approach implied a re-evaluation of social work activities; especially those that concerned consultation with colleagues, analyzing prolematic social circumstances,

and liasing with other social service agencies. When utilizing a unitary perspective, these activities were seen to be at least as important as the traditional casework activities which preoccupied most social workers. It was also suggested that a unitary approach required more work, and a much greater complexity of work, from the social worker. They concluded, therefore, that adoption of a unitary perspective would require that a very strong commitment be made and that a radical review of time-allocation expectations would be undertaken. Group A concluded their presentation with the observation that the workshop had provided each of them with a "mini-experience" of a unitary approach in action. The experience, they said, had moved each member from a simple, intellectual acknowledgement of unitary approaches to the beginning of a personal assimilation of unitary approach principles.

At the conclusion of these group presentations and attendant discussions, it was clear that the day's deliberations had involved all workshop members in looking much more deeply at the ramifications of adopting a unitary approach. Members were moving toward direct consideration of the practical problems involved in instituting a unitary approach within a social service agency or curriculum.

At six o'clock the plenary discussion was rather hurriedly moved into solving the problem of small group reconstitution. The expectation that this would occur had been set forth by the tutors, and most participants had placed their names on the "sign-up sheets", thus indicating

their preferences for work on certain topics.

Group B members, however, stated that they wished to remain together. They believed that the group had reached a point where members would gain much more from meeting again, rather than breaking up into new groupings. They proposed a vague intention to examine some of the implications of a unitary approach to agency organization and suggested that their discussions would probably lead into consideration of the training required to produce a social worker able to adopt a unitary approach. This was a combination of the topics which the new groups would be working on. Three members of other groups said that they would be interested in joining such discussions, but Group B members did not appear to be willing to include them. One "outsider" directly confronted Group B members with their insular attitude but, led by the "artist", the group clung to an almost zenophobic stance. Somewhat arrogantly, one Group B member intimated that his group had "moved further" than any of the others--a claim which, from my observations, had no basis in fact.

The tutors made it clear that they would not force Group B members to separate, so long as they would define their discussion topic specifically enough to offer a real opportunity for relevant and useful enquiries. Once Group B had tried to more substantially specify its topic, the meeting ended. The tutors said that they would post the new group lists after they had analyzed the individual preferences indicated on the sign-up sheets.

One final observation arising from this session may prove useful. Due to job pressures, one member had to leave the workshop on Tuesday and did not return until Wednesday afternoon. He had, therefore, missed all of the Wednesday morning discussions. It was noted that his active participation in this plenary session was much greater than it had been in previous plenary meetings. One felt that there might be truth in the dictum: "A change is as good as a rest." If so, perhaps there are implications for the structuring of such workshops. For example, perhaps two workshop segments separated by a week's "down-time" would encourage a similar reaction in a number of participants. Of course there are many influences which might have contributed to the apparent behavior change on the part of that particular individual; other sorts of research might be more useful in resolving some of the attendant questions.

Wednesday Evening Informal Discussion

That evening a spontaneous discussion arose in the lounge. My experience at previous workshops suggested that discussions of this sort provide valuable insights into participants' perceptions of the instructional sequence; thus, I joined in. A Group D member commented, "I think this is one of the best courses I have even been on." In response, a Group B member expressed a great deal of anger at the "hostility" which she felt emanating from other participants and tutors after her group stated that it did not wish to disband. She believed that she was beginning to

comprehend unitary approaches and felt that a change of groups would have set her, and other like her, back at a time when a deeper understanding was imminent. She very keenly hoped that Group B would move toward resolution on a number of issues and had prepared an agenda proposal for their next meeting; hoping that it would assist them to focus their efforts.

Another Group B member mentioned that, during the afternoon plenary session, she had given a great deal of thought to her reasons for attending the workshop and had considered the extent to which she had acquired new knowledge applicable to her job. She had decided, she said, to leave Group B and join the new subgroup which was set up to consider curriculum planning based on a unitary perspective. That was where her major interests and needs rested. She did, however, find that it had been a particularly difficult decision to reach. She had felt a great deal of cohesive pressure from other Group B members and had experienced some feelings of guilt over the move. A different Group D member seized this opportunity to express his belief that Group B was simply having a "high old time", and that they merely wished to prolong it. This proposition did not correspond with my observations either.

As the discussion terminated, I spoke privately with one of the tutors. He mentioned that one Group B member had reported a feeling that the group had been "caned" for their rebellious behavior. The tutor noted also that the tutors' meeting at the end of the day had been a very

full one. Apparently they had spent a great deal of time discussing the Group B situation. Group B members certainly appeared to feel that they had been attacked and, in drawing up the new group lists, it had been decided that Tutor B would sit in with Group B on Thursday. Having missed the tutors' meeting, while embroiled in the informal lounge discussion, I was not aware of the reasons for this choice. The other tutors, however, had given me the impression that Tutor B was the most experienced--and probably the most expert--group worker amongst them. This may have been a reason for the choice. Later events suggest that it was probably the best decision that could have been taken under the circumstances. Details of the original Group B tutor's feelings about the group are not known.

Thursday Morning Groups

In discussion with the tutors, I decided to postpone observation of Group B until the afternoon session. This decision was taken primarily because: Group B members appeared to be feeling somewhat "put upon"--my presence might have increased the environmental pressures they were feeling, and I had just observed them on Wednesday afternoon and wanted to see what developed in some of the new groupings. Tutor B had no objection to my presence in Group B but I decided to observe Tutor D's new group prior to coffee time, Tutor C's new group in the session after that, and Group B in the afternoon. This left no time for observation of Tutor A's new group but I attempted to tape

record some of their discussion in order to gain some impressions of how they operated and what they discussed.

Tutor D began the new group's first session by establishing a contract with them, as he had done on Monday with Group C. Members had chosen to examine the topic "Organizational Team Implications of a Unitary Approach" and, after a short initial exploration which was ably focused by the tutor, managed to quickly develop a comprehensive agenda for the meeting. For the remainder of the session discussions closely followed the agenda but did not seem to make a great deal of progress toward any resolutions. Throughout the session the high level of intellectual ability and the clarity of verbal communication in the group were impressive. Although "intellectual" may have some negative connotations when one is most concerned with exploring "practical applications", it should be remembered that the workshop as a whole was designed to be an experience in wrestling with relatively abstract concepts. The group was doing exactly that, and doing it well. I was impressed with the idea that if one wished to assist participants in simply learning what a unitary approach was, one would be best to set up tasks that were relatively specific and achievable. In this way, the group members accepted some basic assumptions and began their exploration from that point. If, however, one also wished to encourage participants to explore options for the <u>application</u> of unitary approach concepts to a work setting, one might wish to leave the initial tasks much more open. For, at some point, the

discussion must include exploration of some basic social work assumptions. Tutor D's group, though not progressing very quickly on their "applications-oriented" agenda, were constantly exploring a unitary approach's ramifications for any action system to which it might be applied. This appeared to be a very useful exercise. Before adjourning for coffee, the tutor tried to graphically illustrate his perception of the group's conceptualizations and to set up a specific task for the next session.

After the coffee break, I moved to Tutor C's group and found that they had prepared an agenda which included: (1) "Generalist Specialist Issues", (2) "Conflict/Consensus Views in Relation to a Unitary Approach", and (3) "Integration of Theory and Practice". I was unable to gauge what else they might have done in the early session because they did not appear to have taken up the agenda as yet. It appeared as if they had undergone an experience similar to that in Tutor D's group; i.e., returning to examine basic assumptions. It occurred to me then that there might be a discernable level of disappointment in the workshop by Thursday evening. The groups had set huge tasks for themselves and it seemed highly unlikely that members would feel any great sense of accomplishment from one day's work on tasks of that magnitude

Tutor C's group was attempting to conduct its enquiry through the use of a case study which had been presented by one of the members. Members paid close attention to the points of the case and began to propose possible reasons for the occurrance of the problems exposed. They

did not, however, appear to use any systematic method for analyzing the case. After a few minutes, Tutor C suggested that they start a diagram, in order to explore the case more systematically. He directly asked one member to take the pen and begin. Once the diagram was started, it quickly became the expression of a very complex conceptualization of case circumstances. Tutor C suggested that they "hone" their conceptualization, and asked a direct question to start the discussion once more. Group members responded rapidly and noted a number of important circumstances in the case which had previously gone unnoticed. This was one of the points when I was most impressed with the advisability of using tutors as group leaders. Their understanding of the unitary approach, their experience in working with groups, and their experience in previous workshops on unitary approaches had prepared them to focus group discussions in such a way as to apparently maximize learning on the topic.

On the whole, however, I found that this was a particularly quiescent group. I wondered, for example, whether the atmosphere might have been significantly different had the lights in the room been switched on. Members had been working in a half-light and the discussion seemed to reflect this dullness. Tht tutor occasionally interjected "unitary approach" and "systems" concepts, but group commitment and interest seemed to be on the wane. They did manage to produce some good observations on the case situation but, as one member commented near the end of the session, "I thought we would spend only about a half hour on this."

Tutor C summarized the session, reviewed the tasks ahead of them, and they all went to lunch.

Thursday Afternoon Group

After a free early afternoon, I observed Group B's session. My immediate impression was that they had spent a somewhat ineffectual morning. There were constant statements about having "set the scene" but little was being accomplished. I took quite extensive notes over the next two and one half hours of group-time and have attempted to preserve the flavour of the session by reporting my observations much as they were taken. Again, it must be stressed that these comments are impressionistic. Undoubtedly, my reporting must be biased by my own affiliations within the workshop, but I have tried to present as fair an account as possible. It should also be noted, again, that I refrained from participation in any of the discussions.

- 4:45 - Tutor B tries to get discussion started and Member B (male) follows her lead. He admits that he has difficulties in knowing how to undertake a systematic analysis of a problematic social situation. He asks a question which gives the tutor a legitimate opportunity to introduce more direct input on "systems" concepts.

- Member D (male) has become quite disruptive at times. He tries to subvert the group's work on its self-defined task by changing focus to almost any other topic, and by not participating when his attempted change of focus

is rebuffed. He begins to read while other members are setting out a series of ideas related to task accomplishment.

- Member D's comments give me the impression that he is seriously lacking in theoretical background. He takes an actively "anti-intellectual" stance. In a workshop where so much "intellectual" work is going on, that makes him a disruptor of some note. In combination with Member A (male--the "artist") who is constantly changing focus, and Member B who appears to spend most of his time trying to impress the other group members, group problem-solving becomes a difficult undertaking.

- D appears to be totally involved in reading now. A does not make any contribution for a long time, then suddenly erupts with a provocative statement that has little relevance to the topic at hand.

- I get the impression that Members B, F (female), C (female), and G (male) could work together well if A and D were not present. As it is, frustration builds and very little is accomplished.

- Member B, again, begins to lecture the group. He talks a great deal; usually <u>to</u> people, rather than with them. Again, he appears to be out to impress the others. He is failing to do so.

- 5:00 - Discussion is stuck at a "micro-level". Tutor B redirects to a more "macro-level" topic.

- Now it appears that corporate depression has set in, with Member D encouraging its development. No one says anything of any consequence. Tutor B tries to redirect

the discussion and stimulate some enthusiasm. She asks Member A to graphically represent his understanding of the topic. He goes to the blackboard, starts well, and begins to discover some new avenues for exploration.

- Discussion of Member A's conceptualization develops. However, most participants are unwilling to risk a very full exposition of their own thoughts on the matter at hand.

- Again, Member D presents a plea for an "anti-intellectual" position. He impresses me as a rather rigid sort of individual who very much wants to become, or appear to become, more flexible. His "anti-intellectual" lecture becomes a vehement diatribe against schools of social work.

- Tutor B refers to the conceptualization presented by Member A and encourages further refinement. Member D says he wants to quit. Both the tutor and Member C say that they want D to work with them.

- Member D appears to <u>listen</u> very poorly. He is sulking because the other group members will not do what he wishes them to do.

- The atmosphere is electric and discussion does not stay on any one topic for long.

- 5:15 - Tutor B tries to take the discussion back to the original topic. She writes one of Member G's most relevant suggestions on the blackboard. Some members attend to it, but the discussion continues to leap quickly from one topic to the next.

- Tutor B tries to refocus on the original topic again. "Discussion" is now on three different topics at once.

- Tutor B tries to refocus again. Member D attempts to change the subject but the tutor blocks his effort by suggesting that Member C continue exploration of her last point.

- Member C begins to develop a logical and systematic analysis of the problem area under consideration. Unfortunately, she is very soft-spoken. Member D finds it easy to talk over her. He tries to redirect to an entirely new subject.

- Tutor B returns to Member C's analysis; asking for further clarification and stimulating further consideration.

- There is a lull.

- The tutor encourages Member A to develop the set of thoughts which he had begun to graphically present earlier. A begins to explain his conceptualization, using earlier interactions of group members as an example. He soon finds this leading him into conflict with Member D, and moves on to another example. Member A tends to communicate in analogies, many of which break down rather quickly. He displays a real ability to innovate, but little logical consistency.

- The tutor again encourages him to enunciate his thoughts more clearly.

- Member D begins to tell the group all about what is wrong with social work today. No one seems to be very interested.

- Tutor B goes to the blackboard and tries to illustrate her understanding of Member A's systematization. She

totally ignores Member D's attempts to change the subject. She appears to have concluded that the other members can work without D, but that they cannot work with him.

- Member B calls Member F "Ann", which is not her name. They have worked "together" in this group for three and one half days now.

- I find that much of the interaction makes no sense to me at all. I begin to wish that I more fully understood the participants' motivations for being here. Member D is very aggressive; constantly looking for an argument. Member G tries to act as a "referee".

- Tutor B states that she thinks the group is at a natural point for moving on to another part of the task. She tries to focus the discussion on that topic.

- Members B and D become embroiled in a political argument. Member A uses an element in their argument to refocus on the topic suggested by the tutor. Member F joins him.

- Member D again attempts to change the topic, but Member A will not allow it. A is keenly interested in this topic; it involves "authority", a subject near and dear to his heart. All members except D become involved in the discussion.

- 5:55 - Member D suggests that they bring the discussion to an end. The dinner hour is approaching. Tutor B, however, rejects his suggestion. The group is working better than it has all afternoon.

- Member D does not appear to be listening to the discussion. He says that he is thinking. When the tutor

goes to the blackboard to write down one of Member G's points, D raises a "practical" consideration that has almost nothing to do with the subject at hand.

- At least two unrelated discussions continue at the same time. The tutor and Member F try to reintroduce a unitary perspective on social work practice. However, B rejoins D in a political argument and G becomes involved as well.

- Member F refocuses to discussion of unitary approaches. She refuses to allow D to change the subject. Members B and D, however, begin to argue on yet another topic. They hold the focus of attention for a couple of minutes, then D realizes that it is time for dinner and drops his side of the argument. Everyone is willing to break for dinner.

- After dinner I sit with Members A and D for a chat. I would very much like to understand D's behavior in the group. He tells us that he is ready to quit the workshop. He believes that he had a "major insight" into the unitary approach two days previously. Ever since then, he has felt that he is just wasting time. He also says that he does not feel at all close to "social work people", even though he is employed as a senior social worker. He came to social work from a military and business background and has never felt comfortable in his new occupation. He tells us that he is "basically a working man" and that he is "browned-off with academics".

. . .

- 8:00 - Group B reconvenes and it is immediately decided that there is a need for someone to keep a record of discussion points on the blackboard. Member D says he is willing to take on the job. The group as a whole accepts his offer and gives him a number of points to write down.

- Members seem to have "settled" somewhat. They are working directly on previously set tasks and are staying on topic very well. Part of this change may be due to the realization that time is running short; some may have to do with the fact that Member D is occupied with listening, in order to understand and note down important discussion points.

- Free of D's interruptions, Member B seems better able to contribute to the work. When he is involved in working, he does not try to impress others.

- Member D leaves the blackboard and resumes his seat. He has not been given any points to write down for a minute or two and he wants to talk with Member B. In fact, it appears that B and D have had a chat over dinner. Some sort of conspiracy exists between them now. Member F tries to have it enunciated so they can get back to work on the task.

- As the discussion continues, Member D constantly interjects his perceptions of "the practical world" and slowly takes over the discussion.

- 8:20 - Member B goes to the blackboard and declares that he wants to "go on a flier". He attempts to develop a diagram which represents his view of a basic CQSW curriculum that is based upon a unitary approach to practice.

Other members see a number of flaws in his scheme, but they have trouble finding corrective measures.

- I am impressed with the extent of male dominance in the group. Unless they are willing to be very assertive, the female members do not enter into the discussions at all. At times, some of the males are quite rude about ignoring females' remarks.

- Member B, who describes himself as an "academic", says he is very impressed with Member D's "hard-headedness". At the same time, he is demonstrating increased acceptance of D's rather authoritarian manner.

- 8:30 - Tutor B reminds them that only forty-five minutes of work-time remain. Members B and D attempt to establish a new topic for discussion; slowly, others join in.

- Tutor B encourages the group to begin thinking about its presentation to the plenary session. She suggests that they might attempt to demonstrate some of the benefits derived from their decision to remain together, rather than participating in the reformation of groups (To me, this appears to be presented somewhat "tongue-in-cheek", but members appear to take it seriously). She asks them, "What have you learned out of staying together?" Members report that the greatest benefit has been the development of a "rapport between teachers and practitioners". They have also produced a few unimaginative and rather dreary wall-hangings.

- Member B, the designated spokesman, reviews his intended report to the plenary meeting. It sounds as if he

will do little more than make excuses for their obvious lack of task accomplishment. The group asks him to emphasize that they have been "consolidating" the work of the previous three days.

- The session ends.

. . .

- This group has not been able to do that which it set out to do days ago, i.e., to analyze a problematic social situation in systemic terms. I judge that at least three members would have had a very difficult time if they had joined the other discussion groups and attempted to keep up with current thought in the workshop.

. . .

Midnight "Bullsession"

The following very brief notes were extracted from an informal discussion between five workshop members and myself which took place between midnight and two o'clock a.m. on Friday, the last day of the workshop. With permission from the participants, parts of the discussion were tape-recorded. However, the spatial configuration and acoustics of the room, combined with the quality of the recorder, made for a very poor recording. It has, therefore, only been possible to note down the general impressions of the workshop given by each participant. It would appear to be a worthwhile exercise because the group occurred "naturally" within the environment and members did not appear

to be particularly concerned with censoring their comments for my benefit.

- Female (Group D) - stated that she had known "almost nothing" about unitary approaches before first hearing of the workshop. She felt that the program had "engaged" her, and rated it as "excellent". She talked further about how the workshop had helped her to more deeply integrate a unitary perspective into her view of social work.

- Female (Group B) - expressed resentment over how the Group B decision to remain together had been handled by the tutors. She stated that she had done a good deal of reading in the unitary approach literature over the previous two years and that she felt able to make her own judgements about the discussion topics that would be most useful to her own continued development. Interestingly, a later check disclosed that at the beginning of the workshop, she had reported having read none of the suggested textbooks.

- Male (Group B) - agreed that the Group B situation had been poorly handled. He was, he said, "amazed by the level of childishness amongst professional people". This was obviously a reference to his perception of some of the workshop tutors.

- Female (Group C) - stated that her experience in Group C had not been very rewarding, "because of an unfortunate personality clash." In the reconstituted group on Thursday, things had been much better. She thought that the "interpersonal relationships and dynamics" within Group C had hampered learning; and not only her own.

- Male (Group D) - was generally pleased with the workshop. He believed that it had been a "freeing experience" and that it was significantly better than other courses he had attended because he had no previous acquaintance with any of the participants. Further, he had begun to read up on unitary approaches only a week prior to the workshop and, at instructional onset, had felt entirely unknowledgeable; now he knew much more.

- There was consensus in the group that the NISW application forms, completed by each member prior to the workshop, were very good. They believed that the forms allowed enough scope in response to provide the tutors with sufficient information for judgements on acceptance of applicants. It was intimated that most short instructional programs incorporated highly inadequate application forms.

Friday Plenary Session

Tutor A, the discussion leader for this last session, opened the meeting at 9:30 a.m. by ensuring that all members had been given ample opportunity to view the wall-hangings. She then outlined the proposed agenda for the day and delineated the necessary time limitations to be imposed upon group presentations. Since it was intended that the latter part of the morning be devoted to an "evaluation session", time for discussion would be severely limited.

Group B's presentation further convinced me that they had accomplished very little. Neither were other workshop members very impressed with the group's conceptualization of a

social work curriculum based upon unitary principles. Indeed, there appeared to be little value in any part of the presentation. Their analogy between a systems approach and a kaleidoscope, though colourful, was tenuous. They did, however, make two comments about the workshop which suggested that it had not been a total waste of time. One was that "the lenses on our cameras have become wider-angled." The second was that, within Group B at any rate, they believed that social work educators and practitioners had welded a fruitful interrelationship. Both factions, they said, had accepted a shared responsibility for the development of new approaches to social work training and education.

Tutor D's group was more successful in promoting a stimulating discussion. They had explored "Implications of a Unitary Approach for the Organization of Social Service Agencies". The topic seemed to interest most workshop members and it was evident that the group had worked hard in their exploration.

Tutor C's group reported on its own efforts through the presentation of a case study and discussion. They analyzed the case's most salient characteristics and proposed a rough action plan. During the four Thursday sessions, they appear to have allotted time as follows: (1) definition, clarification, agenda development; (2) exploration of problems presented by a case study; (3) using the case study as a vehicle, exploring the implications of using a unitary approach in the organization of a social service

agency; (4) summary and creation of a presentation plan. The highly involving discussion which followed their presentation was primarily concerned with the question: "Who or what is the client system?" It was obvious that participants could have profitably continued the discussion for some time, but time was running short.

Tutor A's group demonstrated that it had done a great deal of work in designing a curriculum which was based upon unitary principles. Although a number of their conceptualizations seemed somewhat simplistic, generating concern about the amount that had been learned over the week, two highlights stood out. First was the idea of a "macroscope", similar to the "wide-angled" perception mentioned by Group B. It would appear that most workshop members were convinced that their concept of social work practice had widened. Second, they presented a cartoon illustration which ably demonstrated the need for strong links between educators and practitioners in curriculum development and implementation. Interestingly, the groups that examined curriculum planning from a unitary perspective emphasized very similar concerns and proposals.

Although interest was generally high, this pre-coffee session appeared to disappoint many workshop members. Group B members were quick to make negative comments; seemingly out of a desire to belittle the accomplishments of other groups. I began to wonder if there was not a degree of validity to Group B's contention that more might be accomplished by remaining together. As I reviewed the session,

however, it struck me that there had been a number of noticable benefits accruing in the reconstituted groups, directly resulting from new combinations of members and substantial changes in focus. When one tried to assess the progress made in Group B, very little stood out as useful.

After morning coffee, Tutor A reopened the meeting with a request that members put forward any "burning questions" or "pearls of wisdom" arising from their experiences during the week. Turor B led the discussion back to the question raised earlier by Tutor C's group; i.e., "What is a client system?" Tutor A reacted by outlining her own thoughts on the subject, and an excellent discussion ensued. It continued for approximately twenty-five minutes. Tutor A eventually closed that segment of the session by responding to members' requests for "follow-up workshops" and NISW consultancy services. She mentioned the possibility of organizing a follow-up workshop, for a variety of interested people who had attended any one of the unitary approach programs. She also noted that the NISW was willing to provide consultancy services to agencies and other collectives who wished to explore a particular area of interest.

Evaluation Session

The NISW evaluation effort took two forms: a short set of written comments from each participant, and "buzz groups" composed of five to eight workshop members who later presented their comments through oral reporters. The tutors held their own buzz group, and also reported to the plenary meeting. This appeared to offer a relatively simple

evaluation procedure which allowed for a degree of anonymity in the expression of criticism. The only glaring difficulty with the approach stemmed from the fact that buzz group reporters tended to stress the points which they, personally, believed to be most important. While some reporters obviously tried to report very accurately, others were noticably biased by their own opinions. This is, of course, a danger in the approach, and the same sort of criticism may be made against my own reporting. I have attempted, below, to indicate the major comments reported from the buzz groups.

- Comments stated as positive attributes -

- The case material which was provided for small group discussion was useful; especially as a stimulus to early explorations.

- The plenary sessions were a good vehicle for sharing perspectives.

- The reading material sent out to participants prior to workshop attendance provided a good introduction to the subject.

- Tutor A's lecture on the salient features of the various unitary approaches had been very well received and was, perhaps, the most highly appreciated single session.

- Sharing ideas and working with colleagues was seen to be the most valuable and enjoyable aspect of the entire workshop experience.

- The tutors were pleased with both the extent and intensity of commitment to learning displayed by participants.

- Comments stated as negative attributes -

- There appeared to be a general feeling that discussions on the last day had been "diluted". Members had felt a need for a strong concluding session which summarized the workshop's major issues and more firmly "tied together some of the loose ends".

- A number of the social work teachers believed that insufficient time had been allowed for consideration of the curriculum planning implications associated with adopting unitary approaches.

- Some felt that the assigned case material had been focused too heavily on work within social service departments, and that insufficient attention had been paid to other possible areas of application.

- Although the plenary sessions were seen to have been valuable, there was a need for more summarization and interpretation. In particular, it was noted that a closer articulation between subgroup presentations and theory would have been useful.

- The members of Group B, and some of the other participants, felt that the initial "contract" on group dissolution and reconstitution had not been enunciated clearly enough. This had led to a belief on the part of some members, that either the tutors or Group B members had breached the contract. There was mixed opinion as to which side had offended, but most participants appeared to agree that the episode had been disruptive.

- On the whole, most members thought that the workshop had been slightly dominated by the social work teachers.

- Other comments -

- The workshop had been physically exhausting but intellectually stimulating. The physical exhaustion resulted from the large amount of content to be assimilated in such a short period of time.

- The program had undoubtedly provided members with a "wider vision", but its applicability to on-the-job needs remained somewhat doubtful.

- It may not be realistic to expect workshop members to do any preparatory reading prior to attendance. Often they simply cannot find the time.

- During the workshop, there seemed to have been a common assumption that participants had developed a modicum of skill in using a unitary approach prior to attendance. This assumption arose from various long-standing traditions in social work; e.g., the writings of Mary Richmond. Experience in the workshop casts grave doubts upon the validity of that assumption.

- Many participants felt confused at times because of the vast amounts of material covered. Further, becoming accustomed to a "systems perspective" was rather "mind-boggling" for some. Although this confusion might be viewed as a positive condition, because it suggests that participants were extremely involved in pondering the complexities inherent in uniting theory and practice, for

some it might simply indicate that they were lost.

- The workshop certainly had not been boring. It had contained plenty of excitement; and some despair.

Major Issues and Concerns

The information presented in this section is, again, unquestionably impressionistic. Most of it, however, does not result from my direct observation of workshop sessions. Rather, it was generated in the formal interviews held with seven workshop members and my casual discussions with informants during the last two days of the workshop. The series of interviews was conducted late on Thursday and early on Friday, when the workshop was nearing completion. Thus, all interviewees had been exposed to most of the workshop program prior to being interviewed. Interviews were carried our singly and in private. An interview pro forma (see Appendix) was developed on Wednesday evening from my own analysis of the issues and concerns which appeared to most preoccupy participants throughout the workshop. Although I did make an effort to ensure that all thirteen pro forma items received attention in each interview, I seldom found it necessary to ask a direct question. In most cases, the interview seemed to quite naturally cover the content suggested by the pro forma, with little conscious direction from me. Interviewees were chosen for a number of reasons, not the least of which was their availability. Under the stringent time limitations imposed, it was possible to interview only seven members. I therefore chose at least one from each subgroup, and based my choices on information

gained during direct observation of the groups at work. Thus, I chose individuals who appeared to have some special contribution to make. Two were social work teachers who appeared to have excellent analytical skills and a better-than-average grasp on the relevant theory; these two had also done more preparatory reading than most other members. Two were members of Group B who had openly made critical comments about various facets of the workshop; neither appeared to have a strong grasp on the relevant theory and neither had done very much preparatory reading. Two others were employed in the Probation Service, but in almost all other respects were very different from each other: one supervised a Unit of probation officers, appeared to be theoretically naive, had done little preparatory reading, and had been rather forthright in criticising elements in the workshop; the other was neither "teacher" nor "practitioner", she was employed as a fieldwork instructor, appeared to be theoretically sophisticated, had read more of the relevant literature than any other member, and had not vocalized many critical comments about the workshop. The seventh interviewee was distinguished primarily by the fact that she had read none of the relevant literature, appeared to have had special difficulties in her working group, and had stated that she attended because a training officer sent her to gain supervisory skill; her stated reasons for attendance appeared to have no relationship to the content of the workshop. Thus, the group of interviewees was, intentionally, a heterogeneous one.

Casual informants were sometimes chosen because they appeared to be especially knowledgeable about events in which I had a particular interest; or because I came upon them in the midst of discussing some issue of interest. Many other times, conversations came about without any particular intention on my part; I just happened to chat with someone. On a few occasions, workshop participants approached me with the intention of discussing a particular concern or relaying a discrete piece of information. I must have engaged in conversation with each and every workshop participant at least once; many of them spoke with me on numerous occasions. In addition, I very often became involved in conversations with workshop tutors.

In an effort to report upon interviewees' and informants' impressions as accurately as possible, I have made much use of direct quotations. It is almost certain that very few members would agree with all of the statements made. However, unless otherwise noted, I have attempted to choose statements that fairly represent the thoughts and feelings of the majority of members as the workshop neared completion.

Kingswood Hall

During the interviews and randomly throughout the week, members were asked to comment upon the suitability of Kingswood Hall as a workshop centre. Generally, they expressed positive feelings about the facility. No one waxed eloquently over how fine the amenities were, but most appeared to think them adequate. The usual sort of response

was: "It's okay; the rooms were fairly close together, facilities were adequate, we were warm, we were fed . . . I think it's okay." A few respondents commented on poor lighting in the two older meetings rooms--". . . a bit depressing; something to do with colour and lighting. . ."--but this did not seem to bother most participants. The most common criticism concerned a lack of facilities for evening comfort and diversion. For instance, one member said, "I think the areas available for people to informally meet and relax aren't particularly good. That bar and the television lounge aren't really congenial for people to sort of sit around. To go and buy a drink there is almost artificial. It hasn't got the atmosphere of somewhere I would like to have a drink." Occasionally one heard comments about the accommodations being a bit "plain" but only one participant seemed preturbed by them. He said, "I have been spoiled a little bit because I have been to a few halls of residence which are much better than this. I'm a wee bit disappointed by the grotty nature of the accommodation . . . . I don't like standing in the toilet with my feet in water. I don't like being pushed through the wall of the shower by an uncontrollable jet of water." Otherwise, there were very few complaints and some members were very pleased with the College's semi-rural setting.

In my own opinion, Kingswood Hall was adequate. The use of halls of residence for workshop purposes can be a risky business because situations often arise where indigenous residents, institutional fittings and procedures, and

the "intrusion" of other groups are disturbing to workshop participants. Such difficulties must, of course, be balanced against the usually lower costs of such accommodation. In the case of Kingswood Hall, the level of disturbing influence was acceptable and the facilities appeared to be on a par with many throughout Britain.

Half-way through the workshop, an unrelated course of instruction began in the facility. Few participants made any comment about the new arrivals, but a few did suggest that their presence had some influence on the workshop's large-group dynamics. Only one member felt that these influences had been significantly disruptive. In her view, workshop participants had "reacted quite violently" to the other group. She thought it might have been much better if both instructional sequences had started on the same day.

Participant Mix

As we noted earlier, workshop members often expressed pleasure over the fact that attendance had not been limited to only one sort of participant. There was very strong consensus that the membership's diversity in experiential background and agency commitment had proved beneficial to all concerned. In discussing the value of his own small group experience in the workshop, one of the senior social workers said, "I think probably the most decisive factor, though at first it looked like a stumbling block, was having a probation officer in there . . . bringing a very different approach. It made us think outside our closed systems of

education and social work. I think that was very important." He also expressed satisfaction with having had the opportunity to influence, and be influenced by, social work teachers. From the other side, one of the teachers said, "Having practitioners here has helped a lot; especially on the organizational implications of a unitary approach." Comments were made which suggested that there had been a number of instances of antagonism between those who saw themselves as teachers and those who thought of themselves as practitioners. These conflicts, however, appeared to be viewed as natural and expectable occurrances. Members seemed to welcome an opportunity to clarify some of the perceived areas of intraprofessional tension. As one social work teacher said, ". . . of necessity, we have different focuses of interest." It appeared to be generally agreed that the workshop population was slightly overbalanced toward those who viewed themselves as teachers.

There was a more serious criticism about the dimensions of workshop membership. Time and again, participants bemoaned the lack of practitioners experienced in community work. As one member put it, "The unitary approach is stressing 'stop having boundaries between community workers, group workers, and personal social workers', yet--as far as I am aware--there are no community workers present. I think this is a weakness. I don't mind fifty percent of the course being teachers, provided that the other fifty percent is as widely drawn as possible." Many others made similar comments. The very fact that participant diversity was seen

to be a highly beneficial influence suggests that even greater diversity might have proved even more useful. It is not known whether any community workers applied to the workshop; perhaps none did. However, members agreed that such individuals should be encouraged to apply for future workshops.

The following quotation aptly summarizes members' feelings about "participant mix".

> I feel very pleased with it because I think the experiential mix has been good. There has also been a variety of training backgrounds . . . and some members have had a very great deal of experience--as much as twenty years--which I think has been an asset to everyone else . . . and there has been a fairly wide variety of agencies represented. I think we could have done with more people from voluntary organizations or community groups, and I think it would have been helpful to have had somebody who was a community worker within a local authority . . . and somebody who was a community worker outside. But, generally speaking, I think--within the number of people on the course--it has been very good.

Pre-Workshop Reading Material

Many participants commented on difficulties encountered either in locating the suggested literature or, if they were able to find the books, in finding time to read them. For that reason, the papers sent out prior to workshop attendance were regarded as particularly valuable. Members who had read some of the textbooks judged the papers to be a useful refresher. Those who had done very little textbook reading said that they had begun to gain a basic understanding of unitary approaches by perusing the papers.

Only one participant suggested that the pre-workshop material be altered. Her comment, however, articulates

with some of the more general concerns which will be noted in later segments of this report. She felt that, somehow, the workshop required a "sharper focus" and suggested: "I think I would send out different contents in the folder. I would include some diagrams, and I think I would word the summaries rather differently--the summaries of Pincus and Minahan, and Goldstein. This is no criticism of [Tutor A], I'm sure she had to put a lot of work in on them. But I don't think I would have complicated it with all those other things . . . [the other material could have been handed out during the workshop]. I would try to focus people's attention on the Pincus and Minahan and Goldstein frameworks."

Tutor A's Presentation

Tutor A's presentation on Tuesday was warmly accepted and seems to have had a clarifying effect for many workshop members. When asked whether any one session stood out in his mind as being particularly useful, one member responded, "I thought [Tutor A's] presentation was especially good because it outlined a certain amount of material which stimulated our thoughts and directed us into useful discussion. I found that very useful." Many other members, especially the teachers, made similar comments. The didactic presentation apparently gave them something tangible to take home--a set of notes--but, perhaps, also a germinal set of notions with regard to how they might organize and present unitary approach content to their colleagues and

students. The only critical comments about the presentation were that it might have been even more useful if it had been given on the first day of the workshop, and that more time might have been devoted to it. Obviously, participants felt that there was much to be gained from a presentation of this sort.

Case Material

Very little comment was made on the Whittingworth Buildings material. It was familiar to a few participants but none thought that this familiarity in any way reduced its usefulness. With regard to the Rimington case, more critical statements were lodged. There was some feeling that "Rimington" and "Whittingworth" were too similar. One member said, "The discussion material was all centred around social service departments. I think this is a weakness in the course, and I think they ought to have some material produced, say, from a probation officer or a probation office." Certainly, if efforts are going to be made to include an even wider variety of practitioners in subsequent programs, a Study which begins from a substantially different agency focus would be welcomed.

A more common criticism of the Rimington material was that insufficient detail had been included. As one member said, ". . . it was a bit too vague, in that it left too many generalizations about the community at large without providing details. It would be nice if, integrated with that--or even if we had written it ourselves--there was a

survey about the attitudes of the people in this Rimington area. The rest of it was just surmising on our parts. I think that took away a lot from the value. In fact, we didn't even feel that we could surmise what the people felt. A piece of research into what the people felt would have allowed us to do something in view of their needs." The Rimington material contains a large number of "objective facts". If it could be made to appear more real, by the introduction of information on resident attitudes and abilities to cope, it might be an excellent case for a group which was allowed adequate time to handle it--perhaps even a few days. In the time available, and without that additional information, members did not appear to tackle it very effectively.

Participant-Generated Material

On a number of occasions, it was suggested that the most useful case material had been generated from group members' own experiences. For example:

> I couldn't tune in to the thinking very much, and I think it is really only today [Thursday] that it has happened--largely from people presenting their own case material. When we had a question, they have been able to come back with the answers about what actually happened in that case, or what so-and-so would have said. That was one of the frustrations of the case material; you could hypothesize a lot but you never knew what worked and what didn't.

Although this presumed knowledge of consequences may be partly illusory--people often feel compelled to answer when they are asked a question--it does appear that member-generated material seemed more real to participants. Groups were better able to pursue systematic analyses and action

planning when one member had an intimate knowledge of the circumstances under study. Further, it may be that members were more interested in the material because it was related to them as a set of "real-life" events.

> Although we discussed two cases previously, today we discussed a live case that someone brought . . . an ongoing case. And there were a lot of feelings in that--real feelings. Whereas the other two were just . . . I mean, I looked at it as if it was just something made up; it wasn't real. With this new case, I saw how one went about using these new systems. I really enjoyed that.

Groupings

It was earlier noted that the assigned workshop groups were composed relatively evenly on most member characteristics observed. It would appear that the participants were generally pleased with the consequent opportunity to interact closely with individuals from a variety of backgrounds. Those who had complaints about their groups were displeased by "personality clashes", usually with only one other group member. One participant, for example, found her first group experience extremely anxiety-producing and expressed a strong reaction to having had no choice in the assignment. In her own words, her difficulties came about "because of one of my group's members who had been a tutor on my [qualifying] course. I found myself getting dependent upon her, and I could even see her being protective." It is unlikely that a chance occurrance like this will be noticed when groups are composed ahead of time, without participant consultation. Further, it is difficult to predict how individuals might react in such a situation. In

this case, it probably would have been wise for the member in question to arrange, as discreetly as possible, for a change of groups. She reported feeling very tense throughout most of the first three days. A similar experience was reported by another member, who said, "Part of it was a considerable feeling of hostility towards one individual. I just found this person very irritating; and I felt angry with myself that I was letting it . . . and that anger got in my way as well. I don't think I'm the sort of person who lets one individual inhibit me so much." Although such intense feelings were not commonly expressed, one wonders if, somehow, initial group membership might be left less binding for the first day. Then unexplained changes might be more comfortably accommodated.

The transition to a second group on Thursday seems to have been rather difficult for quite a number of members, and was slightly traumatic for the workshop as a whole. Once the recomposition was made, however, many participants appeared to find the change worthwhile. As we will see later, the idea of a transition to "application-oriented" groups was probably not presented clearly enough to workshop members. But, once they began to realize the intentions of the tutors, the new groups appear to have become much more productive and enjoyable. It is difficult to generalize about the experiences gained in those second groups; each of them appeared to be significantly different from the others. Each had its own problems, and each seems to have offered its participants something which they did not gain from their first group.

Questions about what happened with Group B were a major workshop preoccupation on Thursday and Friday. Some of Group B's members expressed dissatisfaction with the expected changes in group composition, saying that it was a "disruption" of the program. In explaining their unwillingness to reconstitute, they suggested that the original course "contract" had been unclear. For instance, one said, "Maybe we were encouraged to believe that we had entered into one contract when we arrived here--in that some of us recall it being said that we could determine whether we would reconstitute or not. In the opening session we were told that it would not matter if we decided to remain in the same groups. Then, two days later, when one group did decide to remain in its original form, we were treated as if we had breached some sort of code. Maybe it was an imagined contract on our part, but the contract was breached and this led to a real rocking of the boat in the middle of the course." Furthermore, Group B members expected that Tutor C would resume his tutorship of them on Thursday and expressed bitterness when they discovered that Tutor B was to take his place. They had asked Tutor C to provide them with a "period of direct teaching on systems theory" on Thursday morning. Whether their perception was accurate or not, they believed that he had intended to return to them but that he was restrained from doing so by the other tutors. Finally, some Group B members were vociferous in stating that by Wednesday evening they had developed a very strong collective will to work, and that they had been at a point

where all of them wished to remain together. "There was an excitement, somehow, that ran through the group at that point. You could feel a gelling; the group felt it was on the verge of an understanding together."

The exact nature of the "initial contract" is somewhat unclear. Early in the workshop, Tutor A did present an expectation that groups would reconstitute on Thursday morning but, perhaps in answer to a question, she also left open the possibility that some group members might wish to remain together. Group B members' perception of Tutor C's willingness to return on Thursday may have been inaccurate. In the last group session on Wednesday, even though asked directly, he made no commitment to return. I must also note that during the Wednesday afternoon session, I observed nothing to indicate that the group was on the verge of any momentous collective understanding. I was not even convinced that all group members wanted to remain together. Certainly they did not all state that predisposition, even when put under considerable social pressure to do so. One member did, in fact, leave the group in order to spend Thursday working on another topic.

Members of other groups were confused about the situation. On the whole, they seemed to feel that if Group B members really thought that staying together would be productive, then they should do so. However, there was also a strong intimation that Group B was ignoring its responsibility for sharing with the rest of the workshop. They remembered that early in the workshop, it was a few Group

B members who had strongly objected to making plenary group presentations. A member of Group C summed up the feeling: "We are a total system, yet that group is apparently finding itself to be quite productive . . . unable to break up and share up. Okay, if they want to do it, it's not an issue that is worth raising here . . . but say this unit was an area team in probation or social services and a group hied themselves off like that. I think we would have to very seriously look at it and say,"'What's going on? You are not playing the game.'" Although Group B members expressed disapprobation with the performance of the tutors in handling the affair, it would appear that most other workshop members did not agree with them.

My own observations lead me to believe that the tensions between Group B and the tutors could have been alleviated, or at least handled in a more satisfactory manner. I was, certainly, never under the impression that Group B had been especially productive, or that it had become a particularly fine learning environment. One or two members hit upon the idea of keeping the group together; in my estimation, at least partly in response to their own feelings of insecurity about reconstituting. They found some support for the idea from one or two other group members. None of these individuals had completed much pre-workshop reading and all were significantly older than the average participant. Furthermore, on the personal information questionnaire administered prior to workshop attendance, three of them had indicated that they had almost no conception of what a

"unitary approach" might be. Their reasons for attending and the benefits they expected to derive from the workshop were entirely unclear; none of them said much at all about why they were attending. I seriously doubt that their intentions bore much relationship to learning about unitary approaches to social work practice. These three, allied with a fourth male, however, weilded a great deal of power in the group. It is my opinion that they unwittingly conspired to protect themselves from scrutiny by refusing to enter new groups on Thursday. I do not believe that any of them was conscious of such a motive. Be that as it may, on Wednesday when they committed themselves to remaining together, Tutor C could have broken up the conspiracy by clearly pointing out that Thursday groups were intended as something different--as opportunities to look at "applications" in a variety of practice settings. If he had made the point that each of them had something special to offer from his own experience, they might well have made different choices. However, once Group B committed itself to continuation, members felt compelled to justify and defend their decision. They attempted to convince all and sundry that they had been having an extremely productive experience. One got the impression that, though they tried, some of Group B's members did not quite believe this themselves. All but one of them, however, did remain true to their decision.

As one might expect, Tutor B's tutorship of the group was not enthusiastically accepted. A few of the members

felt that there was little need for the "clarifying" discussion that took place on Thursday morning. One of their number stated:

> We lost a complete hour-and-a-half session because of that. We were prepared to almost ignore the fact that it was [Tutor B] and not [Tutor C], and we were quite prepared for her to be on the outside of the group . . . until she had assessed the group dynamic. Unfortunately, in my view, she felt that she had to justify the step that the staff had taken; although we quite explicitly said, "We want to go on . . . and you can contribute when we feel it is necessary for you to contribute." We took forty-five minutes of . . . almost recrimination and attempted justification.

This member characterized the tutors' behavior as "immature" and "disappointing". He did not think that they were trying to provide the best possible learning environment under the circumstances. Grudgingly, however, he admitted that Tutor B had been placed in a very difficult set of circumstances and that she had acquitted herself rather well. Other Group B members saw the position somewhat differently and, though they did feel themselves to be in opposition to the tutors, were pleased with Tutor B's efforts. For example, one member who had been staunchly isolationistic and who had described the tutors' behavior as "petulant", said, "Since the opening hostile feelings were expressed, I have had nothing but admiration for the way that [Tutor B] has tackled the thing, . . . and has indeed clarified quite a lost of issues." All in all, he thought that Thursday's sessions had been "very helpful".

The whole affair which arose from Group B's decision to remain intact appeared to be rather confusing for workshop participants. The tutors and Group B members were

seen to be in direct opposition at times. Although very few of them were keenly sympathetic with Group B's position, many members felt that the "initial contract" had not been explicit enough and that tutors would have been wise to more thoroughly and openly discuss the situation at a plenary session.

Outstanding Sessions

Each of the participants interviewed was asked whether any one workshop session had stood out as being especially "good" or "bad". The responses indicate no consensus at all. One said that the first small group session was the best, another said it was the worst. Some voted Thursday the best overall day, while others had nothing good to say about Thursday. The only plenary session which received any mention was the one that included Tutor A's presentation. As we noted earlier, quite a number of members found her presentation to be useful. One respondent stated that it had been the best session of the entire workshop. From the variation in responses, however, it would appear that small group membership and interaction were primary determinants of the quality of experience.

From the welter of judgements, a profile of the "best" sessions emerged. Generally, these were the sessions in which the member perceived his or her small group to be closely knit, mutually sharing, and productive in terms of task accomplishment. It was the small groups at their working best that members appreciated most. The following

account is representative of such comments.

> There was a highlight on the first evening when our group stayed very long over time. We didn't break up until about 10:10 because, with one exception, we all felt very committed to the task as we saw it; and we saw it was producing an end product. We decided we were a very task-oriented group. For some reason which we couldn't fathom out, we were terribly conscientious and we sort of stuck to the task and worked very hard; and we wondered if we were all really goody-goodies. . . . That first evening was a very nice experience, we really shared a lot of hard work.

Interestingly enough, the "bad" sessions were ones in which very little work was done on the tasks at hand. A comment from one member of another group:

> For me personally, I think this morning's [Thursday's] second session was the worst. I think different people have been trying to follow their desired aims, myself included, which has meant that we have not really tackled the tasks which we have set for ourselves.

It could be argued that the adjectives used in stating the question--"good", "bad", "best" and "worst"--were interpreted only in task accomplishment terms, but participants responded in terms of their enjoyment or satisfaction. Obviously, enjoyment and satisfaction went hand-in-hand with shared task accomplishment. Whether either of these bears any relationship to actual learning is another matter.

Program Organization

Although each respondent had a different opinion about how program timing and sequencing might have been beneficially altered, the general opinion appeared to be that sequencing had been very good and that the available time had been structured relatively well. However, all but

one of the interviewees did make some negative comment with regard to the evening sessions. Most felt that the day's work should have been somewhat more "compressed", to allow for more free time in the evening. For example: "I think that by arranging to finish our seminars at 9:15 in the evening, bearing in mind that the setting was away off bus routes, it means we cannot get away from the intensity of the relationships within the groups. I would rather work intensively during the day and finish at dinner." Another respondent suggested that the evenings might have been left free if sessions had been started earlier in the day and the free early afternoons were abolished.

> The afternoons, to me, have seemed a bit of a waste. Rest afternoons . . . rest afternoons for what? I'd much rather have carried on until 6:15 and had the evening free. Also, why start at half past nine in the morning? I mean, most of us are used to starting work at half past eight. Why couldn't we start at 8:30, or even 9 o'clock? Why the half hour for coffee breaks? It would have given us that little extra time. . . . Definitely cut out this quarter past nine in the evening business.

Most other participants said that they did enjoy having the few free daylight hours on two afternoons, but would have preferred working through the afternoons for the sake of having more free time in the evenings.

One unanticipated suggestion was for a scheduled period of private study time; perhaps, half way through the program. In fact, it was the member who had completed the greatest amount of preparatory reading who stated:

> Something that I think I might have enjoyed is some private study time with library facilities . . . or, in particular, a lot of people have had difficulties in getting ahold of the books, and papers have been referred to that we will probably

> still have difficulties in getting. Okay, perhaps I could have asked for them and used some of the free time for that, so it's something to do with my motivation . . . but I think it might have been useful to have something that was actually timetabled. You know, particularly some of the teachers seemed to have got a clearer grasp of it before they came . . . and I think some of the initial problems in the groups had to do with feelings about some people having greater expertise. I know some of it is perhaps fantasy, but if we'd had easy access to remedying some of that gap, it might have helped.

The suggestion appears to be worthy of some consideration. Unfortunately, its institution would probably present problems in gathering together a sufficient amount of relevant literature. It might be a worthwhile course of action if participants were encouraged to bring study material with them to the workshop. Undoubtedly, some members would use the time for a walk in the woods, but others might find more subject-oriented uses for it. As other participants intimated, all workshop members were regularly engaged in full-time employment and often had difficulty in finding the time and motivation necessary to serious study of the literature.

The comment above also relates to the most commonly-stated criticism of the workshop. Almost to a man, respondents felt that there was a serious lack of direct information-giving in the program.

> Although I appreciate that one can learn best in groups, whilst using . . . a tutor as a stimulator--a focuser--I think I would appreciate more explanation of some concepts to begin with; concepts like, "What is a system?" Maybe they don't know the answers, maybe they want those from us. . . . I know things did come out in the discussions, but I think it might have saved time in my group if some definitions, at least, had been given for us to query.

Another member commented:

> I think that in this type of course there is a need for some direct teaching. I don't think that it need be the major part of the course, but I think it would certainly help if there were periods . . . [particularly on systems theory], especially at the beginning. I think it would have saved quite a lot of time. I didn't want them to tell us everything, but I wanted an opening. An opening could have been given, that stopped a lot of movement all over the place; a lot of it not entirely applicable.

Another member perceived the same shortcoming, but suggested that the remedial input might even be presented at a later stage in the workshop.

> It struck me that I had a great advantage, having read Pincus and Minahan, and having wrestled with a lot of this before I came on the course. I wondered what it would be like just to have read the summaries. So, whilst I haven't felt it myself, I have identified the lack. I think that [Tutor A's] lecture on Tuesday was very good, but I think it would have been useful to have had some of the diagrams . . . the systems and functions of social workers . . . projected and explained. I think it would have given people a better framework. Although you could argue that people can, in fact, come to this from first principles--which might be a better learning experience--at some stage, even if it is at a late stage, I think it ought to be presented in that form. It is something to take away with you.

I am emphasizing this point because all but one of the interviewees made it in one way or another, and because I too felt that there was a need for more didactic input. Certainly arguments can be advanced against didacticism, especially in the case of an instructional sequence like this one where participants are expected to learn by "playing around" with concepts and exploring diverse possibilities. But, as yet another participant put it, "The workshop as a whole needed a few more 'reference points', definitions to

react to; in short, more focus."

Task Completion

It was apparent that the tasks assigned to the workshop groups were designed to promote exploration and a more definite understanding of unitary approaches to social work practice. The intention, however, does not appear to have been fully realized; certainly not to the extent that some members felt it might have been. For example, quite a number of participants stated a felt need for greater clarification of the objectives of their assignments.

> I think the major thing would be to get people to understand, from the beginning, that this was not going to be an answer, put down in black and white, to the problems that we face. I think a lot of people came here with the impression that the unitary approach would give them a completely different thing. . . . It has given a wider perspective and an understanding; they will probably go away from it not being as confused as when they first came--but still being uncertain. That wasn't explained to us when we came. . .. Perhaps we thought we would come here and go away with a unitary system, which we would then go back and put into operation. It's not like that at all.

This member wanted it made much more clear that the tasks were intended to promote exploration, but that the participants were likely to leave the workshop with a great deal of uncertainty; uncertainty which would accompany them even as they attempted to apply a unitary approach in their practices.

Even when the tasks themselves were presented clearly, participants were unsure about the higher-level expectations of the tutors. One member stated this directly: "The staff should be very clear about what they are going to expect

from the groups . . . the terms of reference to the groups . . . the contract that was being set." Part of this concern clearly revolved around the reconstitution of groups, and the reasons for that program change. There was a great deal of confusion about the objectives of Thursday's sessions.

> I don't really know whether it's actually been said, "well today we will be concerned with applications." I mean, they may have good reasons for not having said that; but when you think about one group keeping to itself and that sort of thing, it might have been avoided if someone had stated, "On Thursday we are asking you to look from a different perspective and are reforming the groups for that reason." We were told the groups were going to be reformed on the basis of interest, and I may be wrong about their reasons. I just assumed that we are now applying things . . . but, I may be wrong about that.

Another reason for this desire of greater clarification appeared to stem from feelings of inter-group competition. Whether or not tutors intended this reaction to the tasks, it certainly did occur and was a further complicating factor in Group B's isolationism. All members felt an element of competition with other groups, primarily because comparisons betwen groups were bound to be made in the plenary sessions. Further, if there was to be an intended competitive element in the workshop, they wanted a more definite indication of expectations placed upon them. The value of the competition itself, however, was also a question. A number of members expressed their disapproval of it.

> One of the strains came about because we have got to feed back to the plenary group . . . and I think there was a lot of competition as to presentation. Fair enough, one has to share one's knowledge, and

> I think it's good to share the thinking with the whole group, but I think if we had been given more time to relate to our learning--and more idea of what was expected--instead of trying to compete without knowing. . .. To me, and some of the other members, this concentration on trying to compete and present the material was a bit hard-going.
>
> The important thing is not the arriving, the important thing is the journey. It is much more important to work through it than to arrive. I would remove, as far as possible, the competitive element of this feedback thing . . . of the staff making an assessment and saying, "Oh yes, well that wasn't very bad, but of course you missed out on this and that and the other." You got the feeling that you had been a naughty boy, you hadn't remembered all your sums. I would want a more mature approach to it. I would want the recognition of, as I say, the fact that the journey is as important as, if not more important than, the arriving.

To be fair, I did not feel that the competitive element in the workshop was as strong as this comment would suggest. Furthermore, although it is a currently unfashionable viewpoint, one might do well to consider competition as having had some real value for workshop participants. We noted earlier that members had found the most task-oriented groups to be the most enjoyable and satisfying sessions. It might be argued that feelings of inter-group competitoin serve to heighten the drive toward task accomplishment and, therefore, tend to promote greater productivity and more consistent focusing. This argument is part of an ongoing tension for instructors who would use competition to increase motivation and productivity, while attempting to remove extraneous pressures that might detract from the learning experience. Undoubtedly, the optimal point of balance between competitive striving and a non-judgemental atmosphere is difficult to achieve. The point here is

simply that competition may have value for participants if it is not overpowering. Certainly, throughout my observation of the groups, the one that impressed me as being the most productive learning atmosphere was the one in which there was a very strong focus on and drive toward task accomplishment, and ever-present feelings of being in competition with other groups.

Applicability

The extent to which workshop content was applicable to participants' on-the-job needs and opportunities seemed difficult for them to assess. One member pinpointed the core of the problem when she said, "I could probably answer that better in a few months' time." Applicability really cannot be assessed until members have had an opportunity to try out their learnings on real situations. For that reason, I decided to contact members again, four or five months after workshop attendance; to ask the question again. Hopefully, this tactic would give them sufficient time to develop a more informed opinion, without forgetting much of what happened during the workshop. A report on these "follow-up" responses will comprise the latter part of this Chapter. For now, it may be instructive to examine a few of the members' tentative projections into the future at the time of workshop completion.

Those participants who were employed in a social service agency thought that the workshop experience had enabled them to perceive their agencies somewhat differently.

They felt themselves better able to analyze their own working milieu, and to see new means by which their working units might become more effective. On the first point, for instance, a probation officer said, "It clarifies one's thinking. One had vague ideas that there was something wrong . . . that the maximum benefit was not being derived from the system; and now the workshop has put much more structure to it. I can even see it in terms of diagrams . . . in my mind, I can see a structured way now of looking at the system." When asked about the extent to which he saw the probation system as amenable to planned change, he said that, as yet, he had given it little thought. One of the social workers, however, had obviously thought much more deeply and had already developed an embryonic action plan for agency change. She intended to begin by discovering how much her agency's training officer knew about unitary approaches--she suspected it was very little. She intended to encourage him, as well as at least one team leader, to attend a unitary approach workshop. She hoped it would be possible to slowly introduce a unitary approach to her agency's service policies and procedures by building it into in-service training programs.

The probation officer was quite convinced that a unitary approach could be applied in a social work agency. Of social workers, he said, "They have every opportunity for indulging in a unitary approach at a very real level. They have an umbrella of social workers, many have their own special interests . . . they have their community workers,

their group workers, etcetera. They have the opportunity, when a social worker looks at a case, of saying, 'Right, I can deal with this individual, but this individual is part of a system that needs attention . . . we've got the facility, we've got the community here that needs attention . . . somebody has got to go in there over the long-term.' They have the opportunity you see--this is a long-term thing." He was not convinced that the grass was quite as green in the Probation Service, but he was willing to give it a try:

> Within the Probation Service you are much more narrowly focused. I mean, we've got to prepare reports in three weeks, we've got to do the best for a person in one year, etcetera. My probation officers go out into a very troublesome district and they know that things are needed in that area--things that are not necessarily the task of the Probation Service. But there are agencies who ought to be doing it, already set up, and we have to contact them. We have to go to the DHSS . . . and the health visitors . . . and create the kind of working relationship that is not just useful at crisis level.

The social work teachers amongst workshop participants appeared to see the program as having been applicable to their job circumstances. Almost all of them had originally made application to attend, at least partially, because their schools were moving toward teaching an "integrated perspective" on practice. Speaking of the benefits which he had derived from the workshop, one teacher said, "I think it has given me a much wider material base to base my work on. I'm supposed to be developing a teaching approach using this method. At least it has helped me to sharpen up my theoretical understanding. It has also given me material

. . . I think I could devise my own material now; my own case studies."

In the comments on applicability, an interesting side-issue was raised by two members. Both suggested that their learnings from the workshop were probably all the more applicable to their work because a colleague from the same agency or school had also attended the workshop. They felt that this would make it easier for them to initiate and sustain interest in unitary approaches on the job. Two participants could be mutually supportive and doubly influential. That belief led, later, to a suggestion that the membership of future workshops include a number of people who work within the same agency or geographical area.

Major Deficiencies

Participants intimated that there had been two major areas of deficiency in the workshop. Both of these have already been mentioned in passing but deserve greater attention. The first was a commonly-felt need for greater clarity and specificity--more "focus". Undoubtedly, some of this feeling resulted from the nature of the workshop content itself. "Unitary Approaches" are not, in themselves, all that unitary; or all that well conceptually developed. They are not highly discrete entities which one can unequivocably describe and explain. However, criticism went further than that. Participants wished for a clearer presentation of the terminology and concepts of "systems theory" as well. They wanted: ". . . a core session which provided better explanation of what a unitary approach is and how it

relates with 'systems thinking.'" Along with that, they felt the need for a clearly stated summary at the end of the workshop--"something to take away with us." It appeared to be generally felt that the staff really had not provided enough direction. At the same time, more consultative interaction was desired. As one participant said:

> I mean, they have acted as a resource group; given information and set tasks, etcetera. But I haven't felt that--it's hard to identify. I've seen them very much in role I think. I wanted them to play a more active, participatory role and make more joint decision . . a little more staff-student interchange about how people wanted to organize their time and learn.

So, on the one hand, members said that they wanted a somewhat "neater package"-- more explanatory, direct teaching input--at the same time, they would have enjoyed a closer dialogue with tutors regarding the ways in which the workshop was progressing and how it might be beneficially altered in process. These desires were not seen to be contradictory since the first dealt, primarily, with content while the second dealt, primarily, with process.

The second major deficiency stated concerned the variety of participants present. As noted earlier, members enjoyed their collective diversity and thought that it was a very useful characteristic of the workshop. However, they desired even more diversity: ". . . a wider range of agencies and, perhaps, a few more people in top management positions . . . definitely some community workers." One member had been told that previous workshops had included teachers only. He felt that such a workshop could not have been of much benefit to its participants.

Benefits

Contrary to the impression that might have been given thus far, seriously critical statements about the workshop were really quite rare. Even one of the most vociferous critics eventually said, "I think I have come to gain a reputation of grumbling about the course, but basically I'm fairly happy with it." During the interviews, an effort was made to determine what, if anything, participants felt they had gained from the workshop. Their responses were found to have a number of common characteristics, and are summarized under the headings below.

- General Comments -

Each person who tried to make a general statement about what he or she saw as the overall benefits of attending the workshop, presented a somewhat different viewpoint. Most of them, however, felt that they had more fully integrated, or come to a stronger commitment about, unitary approaches to practice. The following three quotations illustrate the range of these responses.

> I think it is something about feeling more emotionally committed to the integrated approach than . . . than just intellectually. I mean, I had a limited degree of intellectual commitment, but I think I have gone a lot further than that now. I'm beginning to think in systemic terms.

> . . . an appreciation that a concept of integrated methods is a lot more complex than I originally fully appreciated. It made me question--or reinforced and highlighted my previous questioning, really--'What is social work all about?' You've got to ask more wider questions, and [the workshop] throws up a series of these that are very difficult to answer. It boils back, eventually, to values and responsibilities; and where do you assess these?

> . . . experience of putting these ideas into some sort of practice. I had heard about them but I hadn't, in any experiental way, juggled around with them. . . . I think, for me, it's hearing what other people feel, coming on it in an oblique way, and then sort of trying it to the task.

The other set of overall benefits reported were even more directly concerned with the value of colleague interaction. One member, for instance, said, ". . . the course has helped me to conceptualize about the need for interagency linkages, largely through contact with people from other work settings." A social work practitioner thought that his contact with teachers had been extremely valuable:

> Looking back, my original ideas of a unitary approach were very vague indeed; and the teachers in the group, picking up practitioners' ideas, developed them into a kind of theoretical context which made much more sense. It had to do with something that I hadn't had the opportunity of doing for a long, long time . . . of coming into close contact, as a practitioner, with social work teachers.

- Benefits to Agency Personnel -

Those who were employed by direct service agencies said that the workshop was particularly useful in providing a means by which they could begin to analyze problematic situations, including the functions of the agency itself. They emphasized that it had led them to a greater realization of what their own agencies were doing, and that it had highlighted some of the difficulties in breaking down dysfunctional boundaries between both agencies and specialisms. When asked what benefits he had derived from the workshop, one practitioner said, "A wider perspective . . . to get one step back from the agency involvement and see the role and

function of other agencies . . . beginning to feel less threatened by other agencies, less protective of my own agency's function."

- Benefits to Teachers -

The special benefits stressed by social work teachers were two-fold. The first had to do with development of a deeper understanding of unitary approaches in social work. This was seen as crucial for the teacher because he had to explain these approaches to his students.

> I have now had the experience of applying what before was an intellectual concept; something I'd got out of books. I mean, we've only been here a few days but I've been able to do mental acrobatics with it. I'm easier with the theory now. 'Integrated methods', which is what I call them, were totally new to me and I've had to devise a syllabus based on this . . . and my own practice was miles away . . . I think now it will be easier . . . [when answering students' questions] . . . to think 'Now what is the unitary approach to this?'

Another commented that the workshop had changed her attitude toward teaching. "I think," she said, "that I've understood more clearly the place that there is for specialisms."

The second special benefit cited by teachers was a gain in information about approaches to, and materials for, the actual teaching of unitary approaches. It was previously noted that one teacher believed he had become much more able in devising his own teaching materials. Another was not even all that concerned about teaching from a unitary perspective; she was simply pleased that the workshop had ". . . given me a look at how you teach, how you run groups . . . which is not one of the stated [workshop] objectives;

it's just that it is something I am constantly trying to learn about. You get very little help in attaining teaching skills, and so you have to pick them up wherever you can. This has been one of those experiences where I have obtained a few ideas."

In summary, although some critical statements were levelled against the workshop; participants generally felt that their participation had benefited them. Further, when asked directly, most intimated that they would have made very few changes in the program had they been the workshop's senior tutor.

Objectives and Learning

Early in this report, we noted that tutors' and members' sets of intentions/expectations appeared to be essentially congruent. The primary emphasis in both appeared to be upon knowledge acquisition to promote understanding, and the application of that understanding to social work practice--whether it be in direct work with social work clients or in teaching. It is apparent from their responses at the end of the workshop that many members believed these major intentions/expectations to have been met, at least partially. Just as there were other, more minor, expectations stated, there were other, more minor "outcomes" observed. One might, in fact, wonder whether the achievement of other "outcomes" (e.g., increased understanding of small group dynamics) detracted from the workshop's efficiency in reaching its acknowledged objectives. However, these other effects appear to be enormously outweighed by the strength

of the major benefits cited. The majority of respondents expressed a strongly-held belief that the program had helped them to develop a deeper understanding of unitary approaches. At the same time, one must not forget that the results reflect only members' expressions of belief. There is no evidence to indicate that, in this particular instance, members' beliefs positively correlate with the actual extent of learning. No more "objective" tests were undertaken to assess actual learning. Furthermore, members had a great deal of difficulty in assessing the applicability of their newly-found knowledge. One of my foremost goals in using a follow-up questionnaire was to gauge the extent to which members perceived their workshop learnings to be applicable in their working milieux.

Follow-up Questionnaire

In mid-June, 1976, approximately five months after the workshop, a "follow-up" evaluation form was mailed out to each participant. It asked the four questions listed below and requested that respondents reply in any way they desired. The four questions were:

(1) "Looking back on it, was the workshop a worthwhile experience? If so, in what ways?"

(2) "Over the past four months, have you been able to put into practice any of the knowledge or skills you acquired at the workshop? If so, please give an account of what has happened and how your workshop experience assisted you."

(3) "To what extent do you think the unitary approach is realistically applicable to your work? What are the major constraints on its use in your particular situation?"

(4) Can you suggest ways by which the workshop might be altered in order to provide a more useful experience for future participants?"

Of the twenty-seven possible respondents, eighteen completed and returned the instrument. All eighteen responded to each of the four questions. Although this number of responses represents a fairly good return rate (70%) on a questionnaire mailed out five months after program completion, there was, of course, a possibility that those responses were in some way biased by factors which might have been related to respondents' proclivity to receiving, completing and returning the form. For this reason, a rough check on the statistical representativeness of the respondent "sample" was carried out on four characteristics: age, sex, job, and small group membership. The average age of respondents was almost exactly that of the total workshop membership. While the total workshop population contained fourteen males and thirteen females, the respondent group contained ten males and eight females. In terms of small-group membership: Groups A and B were well represented (5/6 and 6/7 respectively), while C and D were not so well represented (3/7 and 4/7 respectively). Figure 5-3 indicates the job categories of respondents as compared with those of the total workshop population.

FIGURE 5-3

JOB CATEGORIES OF RESPONDENTS AND TOTAL WORKSHOP GROUP

Job Category	Respondents	Workshop Total
Social Work Teachers	8	9
Teacher-administrators	2	4
Direct Practitioners	1	2
Fieldwork Instructors	1	2
Agency Administrators	1	1
Seniors or Supervisors	5	7
Supervisor/Training Officer	0	1
Training Officer	0	1
Total	18	27

Both teachers and supervisors were well represented, and at least one representative from each category, with the exceptions of the training officer and the supervisor/training officer, responded. It should also be noted that two of the four Probation Service personnel responded. With regard to all four characteristics, the proportions of the respondent group were not identical to the proportions of the total workshop group--that could have come about only with a one hundred percent response. However, the proportions approximate each other quite closely. If there is any bias, it would appear to be only slightly in favour of the members who participated in Groups A and B. A summary

of responses to the four questions is presented below.

Question #1 ("worthwhile experience")

All respondents stated that the workshop had been a worthwhile experience for them. The most common response--mentioned by over one half of respondents--suggested that the opportunity for sharing viewpoints, ideas and experiences with colleagues was the major perceived benefit. For example, the fieldwork instructor said, "I found the mix of practitioners from a variety of agencies and from social work courses, invaluable." One of the supervisors put it slightly differently: "Interdisciplinary intercourse forms the basis for a unitary approach. The workshop provided practical evidence of this." Members who were employed as social work instructors presented this viewpoint particularly strongly; one, in fact, called it "reality-testing through contact with the field", while others suggested a wide variety of reasons for the high value they placed on the workshop's sharing element. Witness the teacher who reported a gain in self-confidence because he learned that "other, more experienced teachers were no more secure in their theoretical understanding than myself"! It is unlikely that such feelings would have been so strongly expressed had the workshop program not depended so heavily on the small group discussion format. It was that format which encouraged, in fact demanded, a high level of interpersonal involvement. Had the program been organized along more traditional didactic lines, the opportunities for interpersonal sharing would have been much more limited.

The second most common response to Question #1 had to do with respondents' feelings that they had "stretched", "re-ordered", or "clarified" their understandings of unitary approaches. A few saw the workshop as merely a good introduction to unitary approach concepts. However, a larger number--perhaps with different theoretical backgrounds--noted that the program had helped them to understand relevant concepts more thoroughly and re-order their previously held formulations; that it had altered their overall view of social work and increased their awareness of alternative practice modalities. So there was not simply a feeling of having gained scraps of knowledge about unitary approaches to practice, but also a strong indication that the acquisition of knowledge had begun to produce some changes in the conceptual tools used and the values held. A number of respondents used the word "assimilation" to indicate, perhaps, the development of a more deeply integrated system of knowledge and values.

Responses which fit comfortably within these two categories made up the bulk of replies to Question #1. There were, however, a number of other responses which are interesting and, quite possibly, instructive. Six or seven respondents mentioned that they had enjoyed and/or found valuable the opportunity to retreat from their normal work and domestic responsibilities. A few of these comments suggested that this brief respite had given them an opportunity to concentrate their thoughts on unitary approaches, their implications and applications. Most, however, were

simply expressions of gratitude for the opportunity to "get away from it all" for a few days. This is a relatively common response to questions about the value of workshops, short courses and conferences. Although it is often accompanied by expressions of guilt, it might be viewed as indicating an entirely legitimate secondary function of training and development activities.

A number of other respondents concentrated most heavily on expressing appreciation of the opportunity to study possible applications of unitary approaches; particularly applications to the respondent's own work milieu. Further, quite a number of workshop members expressed pleasure with their involvement in the cooperative, task-oriented work that occurred in the small groups. For some this was a simple expression of enjoyment, but others noted that the experience had benefited them in developing abilities to perform better on the job. For instance, one teacher stated that he had found the workshop to be a "valuable experience in the use of small groups as learning vehicles."

Finally, there was one respondent, from the Probation Service, who made no mention of any benefit or enjoyment accruing to him. However, he believed that the workshop had been valuable for others. He said, "I feel my participation was worthwhile mainly in helping those from Social Services, the lecturers from Polytechnic, etc., to clarify their thoughts. This is not meant in any way to imply a 'know-it-all' attitude, but to reinforce my comments that the Probation Service is much nearer a unitary approach

than many other agencies." From discussion in the workshop, it seems likely that many participants would question the validity of his assessment of the Probation Service. The interesting point, however, is that five months after the workshop, this particular member reported absolutely no benefit to himself. On the last day of the workshop, he had indicated his enjoyment of the program and announced that it had taught him how to "think more systemically and diagram out situations".

Question #2 (use of "knowledge or skills")

Since teachers made up almost half of the respondents, the most common response to Question #2 was that participation in the workshop had, in one way or another, assisted in the development of course curricula based on unitary approaches. Almost every teacher-respondent noted this area of application. For example: "Have introduced limited version of unitary/integrated approaches into social work methods sequence in first year. It is hoped next year to tie this in with a broad micro/macro introduction to social work and with integrated methods fieldwork placements." We have noted that many of the teachers attended the workshop with the very intention of preparing themselves for an expected curriculum change toward unitary approaches, even to help create that change. From the follow-up responses, then, one would judge the workshop to have been quite successful in assisting them to realize their preparatory intentions. They arrived wanting knowledge, to prepare themselves for

expected curriculum changes. Five months later they reported that the workshop had made them more knowledgeable and that the appropriate changes in curricula were being instituted.

A second response, closely associated with the first and again from the teachers, was that the workshop had provided something of a model; an example of some of the means by which one might teach a unitary approach. Quite a number of the teachers reported that they had used some of the teaching methods and materials from the workshop in their own classes. For example, two participants who teach on the same CQSW course have: "taken a leaf from the workshop methods book. That is, we have devised some participatory exercises for the students; we have also tended to break down our lecture group (36 students) into subgroups after the formal teaching." Others said that they had used the workshop's case studies in their own classrooms. One used "some of the visual diagrams that the small groups produced as illustrations for teaching. . .." There appeared to be consensus that the use of workshop methods and materials had helped to clarify presentations to students, and to produce better teaching from, and of, a unitary perspective.

Members who were senior social workers or supervisors also reported that their workshop experience had proven to be useful on the job; particularly as an aid to the development of their social work teams. One, for example, believed that in communicating his new knowledge to team members, he had helped the team to "more clearly conceptualize its working methods", and that this increased clarity was proving to be particularly useful when more than one team member was

involved in the same action system. Another said that his new knowledge had assisted him in ". . . improving team management--towards more democratic team self-management." The results, he said, can be seen not only in the team's closer definition of tasks, but also in clearer contracts between the team and its clients. Furthermore, the team's relationships with other agencies have also improved. For example:

> In the past, referrals from doctors, health visitors, etc. were always dealt with but we never notified the results of our involvement to the person or agency who made the initial referral. Now I have instituted a system of notifying the person concerned and this has benefited our relationships with other agencies.

Another supervisor noted that increased emphasis on a unitary approach to practice had been particularly useful to him and his staff in reviewing their relationship to, and use of, volunteers. He believed that their volunteer programs had improved considerably.

In order to accurately report on responses to Question #2, it should also be mentioned that one supervisor thought the workshop experience had very little value for his work. Another, though believing that the workshop was useful to him personally, found that he was encountering a great deal of agency resistance to unitary approaches. In summary, he said, "I am continuing to seek ways of introducing this theory to the team and the department. However, no real progress has been made as yet." He did, however, believe that the workshop had rendered him much more able in communicating about unitary approaches to practice.

There was one other response to this question which appeared more than once. It had to do with yet another "spin-off effect". Some of the teachers noted that they had observed their work associates developing greater interest in, and understanding of, unitary approaches. Many of these were colleagues who had not undergone any formal instruction which might account for their increased interest and understanding. Although there are probably numerous reasons for this phenomenon, not the least of which might be changes in school policy, it is likely that some of the interest and understanding has been generated by contact with the workshop participant. Two, in fact, had organized "mini-workshops" for colleagues. They reported that colleagues were actively seeking further knowledge and introducing ". . . more integrated teaching sequences across disciplines."

Finally, two respondents felt that they had not put into practice any knowledge or skills derived from the workshop experience. One did not believe that he had acquired any new skills and did not think that he had really applied any of the knowledge which had been gained. The other, who had just changed jobs, did not feel able to apply unitary approach principles to his new situation. Both of these respondents were working in administrative capacities and did not see the lessons of the workshop as being readily applicable to their work. The second one, in particular, stated that a unitary approach might be useful for "case analysis", but not for agency administration and policy

development. He said, "I have been unable to apply the principles we discussed since at this particular time I am not involved in 'cases'." The occasional comment like this makes one wonder about the extent to which deeper understanding of unitary approaches actually did occur!

Question #3 ("major constraints")

A few respondents firmly stated a belief that unitary approaches were "generally realistic" or "realistically applicable". However, most made no such overall judgement. It did appear that most respondents had found a relatively high degree of applicability to their particular jobs; even if they had observed a number of possible, and sometimes very practical constraints. The senior social workers were especially concerned about difficulties associated with introducing and gaining acceptance for a unitary approach in an agency which strongly holds a traditional casework orientation. These difficulties were not well defined but included: a generalized feeling of institutional resistance to change; the recognition that few workers have had any training which might dispose them to adopting a unitary perspective; and a belief, common amongst caseworkers, that a unitary approach would demand of them a greater range of skill than they would be able to demonstrate.

At the same time, there were suggestions that unitary approaches, as theoretical constructs, contained inherent difficulties which reduce their applicability. Some, for instance, suggested that unitary approaches are

presented at such a high level of abstraction that applications become rather obscure. Furthermore the level of abstraction was seen to pose a special difficulty for workers and students with little theoretical social science background. Others observed that the successful use of a unitary approach required a ". . . well-integrated team, capable of honest sharing of strength, weakness and talents . . .." Since creating such a team consumed a very great deal of time and effort, this constraint was seen as especially inhibiting in a situation where staff-turnover was high. Even in a well-integrated, highly committed team, the unitary approach, they said, demands a high level of team consensus which can only be produced through the expenditure of large amounts of valuable time and effort. Further, although the approach was seen to be immediately and realistically applicable as a tool in assessment, there were suggestions that it is a rather static model--requiring modification to take greater account of process and change, and in need of much clearer definitions as regards the skills implicitly required for utilization.

The teachers stated that a major constraint was difficulty in providing students with fieldwork placements that give them opportunities for using a unitary approach to practice. This problem was seen to have a number of facets: (1) the general orientation toward practice within most agencies is not very compatible with unitary approaches; (2) supervisors in such agencies require careful preparation before they are ready to take students who will be learning

to operate from a unitary perspective; and (3) a unitary approach appears to imply and require the existence of management styles which are not currently found in many agencies. As one respondent put it: . . . many placements are not able to stimulate students to operate/think from a unitary perspective, hence the opportunity to test out new ideas depends very much upon student initiative which is often rather limited."

In association with these fieldwork-related problems were a number of felt difficulties arising from the "newness" of unitary approaches. It was suggested, for example, that there are difficulties in helping fieldwork instructors to acquire the new understandings with which students will confront them. This was seen to be exacerbated by the fact that fieldwork instructors must be consulted and planned with at the same time as they are floundering about, learning the new approach. Furthermore, implementation of a unitary approach in a curriculum requires that a great deal of time and effort be spent on discussions amongst teachers, administrators and fieldwork instructors. In other words, a number of tasks must be carried out concurrently, and--to a large extent--success in one depends upon accomplishments in another. In addition, existing staff are reported to often have very strong value commitments to traditional social work approaches, and "service teachers" (from other departments), who have no social work experience, must also be integrated into any system which attempts to use a unitary approach. Much time and effort must be spent on preparing both sorts

of individual before integration is possible. If we add to this problem complex a number of perceived difficulties stemming from questions like: "How much can we teach in a limited period of time?" and "How do we go about setting priorities for what we teach (generalist vs. specialist issues)?", the difficulties may seem somewhat overwhelming. It is little wonder, then, that teachers were apparently trying to pace the introduction of unitary approaches into their course sequences. As one said, "I think it will take ages to sort ourselves out."

As we noted earlier, one representative of the Probation Service was convinced that his organization was already well on the road to using a unitary approach. At the same time, he remained rather dubious about the advisability of such an undertaking because, to him, it smacked of ". . . community involvement and similar which is so often closely associated with conflict." As he saw it, conflict is damaging to the Probation Service and to the community as a whole. Another Probation Service representative, however, did not see any value in that point of view--though she was concerned that implementation of a unitary approach would be very difficult precisely <u>because</u> that viewpoint was prevalent in the Service. She was also concerned that current training in the Probation Service was so heavily based on casework theory. She went on to say that there were a few legislative constraints, such as limitations on the use of certain facilities or programs, but that she was optimistic. I feel", she said, "that many of the objections to the application of

a unitary approach in the Probation Service are defensive. Accountability to the courts in respect of individual clients can still be maintained. Most courts will accept imaginative work if the reasoning for it is carefully spelled out."

Two of the most administration-oriented respondents foresaw constraints upon using a unitary approach because it was stated in too abstract a manner; it was not seen as "practical" enough. The senior administrator said, "The only constraints on its use are the tendencies for some social workers to think that concentrating on a particular approach is more rewarding than using all available and appropriate approaches." A teacher/administrator summed up many difficulties when he said, ". . . the major constraint is the lack of trust between varying disciplines."

Question #4 ("alterations")

On the whole, even the responses to this question implied an overall contentment with the workshop. Generally positive comments included statements to the effect that the advantages in the workshop program far outweighed any in alternative programs which the respondent might be able to propose. One of the teacher/administrators, for example, said, "I'm not sure any different structure would have given me a more useful experience." He was especially pleased with the use of task-oriented small groups. So were many others. No respondent suggested a comprehensive alternative to the existing workshop program. There were, however, a number of more elementary criticisms; criticisms which might suggest that certain smaller changes in the workshop would prove

beneficial.

As might be expected from earlier comments, a number of respondents suggested that there was a need for larger amounts of direct input of theoretical material on unitary approaches and systems theory. As one of the teachers put it: "All participants arrived with some knowledge of systems theory but there was an apparently great variation amongst them in that regard, and everyone floundered a bit from time to time. There was an obvious need for more grounding in that area." Other respondents made similar suggestions. One, for instance, felt that more time should have been devoted to Tutor A's presentation, and that it might have been less "confused". Another said that ". . . basic introduction to the elements of the theory would have more usefully come at the beginning, rather than the middle, of the workshop."

Like much of modern social work instruction, the workshop program was based upon principles of heuristic learning. That is, the instructors tended to withhold direct definitional and theoretical statements, while encouraging members to discover and develop their own relevant conceptualizations. To the extent that they wished to encourage exploration of basic theoretical elements and "active", as opposed to "passive", learning behavior, this was probably a wise policy. However, the policy would appear to have had at least one rather unfortunate consequence. By substantially ignoring the possible benefits to be had from more didactic inputs, the tutors apparently left many workshop

participants with an inadequate definitional component upon which they might have begun their explorations. It would appear that many members arrived at the workshop with virtually no conceptualization of unitary approaches. Thus, many did not possess even an adequate conception of the boundaries to the area being explored. It is suggested, then, that tutors might have found profit in taking a somewhat more expert stance and using more didactic teaching methods early in the workshop. Should more time be available for future workshops, the "muddle through" approach may prove entirely adequate. If, however, time limitations remain stringent, tutors would do well to consider alternative means for the presentation of definitional material early in the program.

The most commonly-made suggestion was for a follow-up course, held some months after the workshop. These responses revolved around the notion that a follow-up experience would allow participants to more realistically assess and plan applications of unitary approaches in their working circumstances. As one member said, "I feel ready . . . to make much greater strides . . . if we had another week now."

There was also a considerable amount of comment concerning the composition of the workshop membership. Both teachers and others suggested that there had been an overly-strong representation of teachers. Some said that, because of this overloading, the views of social work practitioners--especially with regard to the content of social work training courses--were not given enough voice. Others wished that

there had been a few community workers and community work instructors present. Another would have appreciated contact with ". . . more experienced fieldwork teachers--particularly some who are associated with my own teaching programme." Yet another suggested that the workshop should have included Educational Welfare Officers and, even more importantly, in-service training officers. She posited that the training officer was in an excellent position to influence agency change toward a unitary approach by filtering information into the organization and encouraging appropriate training programs.

Again, there were a number of comments suggesting that, somehow or other, more "practical" input was needed; a "practical demonstration" or "live example of the approach in action". For instance: ". . . it would increase the credibility of the approach if some practitioners experienced in its application had been included among the NISW staff. This might also help those who found the workshop too academic."

Three comments had to do with the way in which the small groups were composed and maintained. Obviously, all three arose from a concern about the "confrontation" which supposedly took place. There was, however, no agreement as to what that "confrontation" was about. One suggested that "a greater flexibility" with regard to changing the organization of the workshop in process would have produced more harmonious relations. As an example, the respondent suggested that greater attention to group dynamics on the part of the

tutors ". . . would have avoided the confrontation." Another thought that ". . . a more specific contract, linked to tasks and group functions ". . . would have helped." The last said, "Perhaps keep lecturers and fieldworkers mixed throughout the workshop to avoid the unfortunate split which occurred between the two. Or else engineer the thing to take account of this conflict and hold some sessions for the opponents to have it out frankly, as a start to working together." Obviously, all three respondents noticed some tensions within the workshop membership, but they did not agree amongst themselves about the causes, properties, or remedial actions required to alleviate them.

We have noted that participants found benefit in being able to meet and share thinking with their colleagues. One respondent to the follow-up questionnaire suggested that this aspect of the workshop should be strengthened by further encouragement of activities which provide opportunities for informal sharing. He thought that, "The success of this workshop was . . . helped by the bar, games available, etc., but I wondered how far this was due to design and how much was accidental."

"Timing and sequencing" concerns were mentioned only twice: once by a member who would have preferred to work in the afternoon rather than in the evening, and once by a participant who had found the small group sessions so much more useful than the plenary sessions that he was prepared to advocate a reduction in the allotment of time to plenary sessions.

Finally, a representative of the Probation Service suggested that the program should not be ". . . obviously structured to the needs of social workers and lecturers." He hoped that future workshops would ". . . use some probation-oriented material, including court reports and after-care cases."

Summary of Major Suggestions

The suggestions presented in this Section result from sifting through all of the material gathered. All of them have, to some extent, been discussed in earlier Sections. All were made, in one form or another, by a number of workshop participants and tally closely with conclusions drawn from my own direct observation of the workshop.

(1) A large number of workshop members hope that some arrangement will be made for at least one "follow-up" session. Generally, this was presented not merely as a re-gathering of workshop participants, but as an opportunity to meet with others who had attended the various unitary approach workshops in order to discuss the experience of attempting to implement a unitary approach--whether in an agency or in a school. One workshop member, in fact, demonstrated the value of a follow-up session. For her, the workshop itself was something of a follow-up to a weekend workshop on unitary approaches which she had attended only weeks before. She said, "I participated in this weekend workshop which was brought together for fieldwork teachers, which [Tutor A] had been present at . . . and, while what she said

had a lot of impact on me then, I've been able to link up a lot of what has happened this week . . . link it back."

(2) In one way or another, a live demonstration of the unitary approach in practice would probably prove to be very useful. Participants suggested that they would have appreciated direct contact with someone who had used a unitary approach extensively and could relate his or her own experiences with it.

> It would have been helpful to have participating in this workshop . . . [someone] who has actually had a living experience . . . [of using a unitary approach]. Now whether you want that from a staff member or you need a few practitioners who have used it, I don't know. Basically, I would like to talk with someone who can describe the unitary approach working, or not working, in his particular situation.

(3) At least one session of direct explanation of "systems" concepts and terminology, preferably early in the workshop, would give participants a clearer point of reference from which to begin their explorations.

(4) Some sort of summarization of the workshop, coupled with exposition of a few of the most important principles underpinning unitary approaches, should be given near the end of the workshop. This would assist participants in leaving with some more integrated conceptualization of unitary approaches; a slightly more "gelled" set of ideas.

(5) Some effort should be made to allow for more free time in the evenings. This might be accomplished by starting earlier in the mornings, and cutting back on some of the free time offered during the afternoons.

(6) An even greater diversity of participants should be encouraged. Workshop members were particularly keen on the inclusion of a few community workers because it was felt that they would be able to make especially relevant and valuable contributions.

(7) Closer student-staff consultation over staff intentions, student desires, and workshop processes could be helpful to all concerned. As the following quotation suggests, the problems encountered in the transition to new small groups on Thursday might have been partially eased by greater explicitness on the part of tutors.

> Reconstituting the groups seems to have been a big issue. . . . I think that if I were running the course, I would make it more explicit that these were application groups, and that they were different from the groups we had before. But I think the other unacknowledged caution--or it may not be acknowledged by the staff--is the weaning process . . . because there is always this separation thing with courses. Perhaps one way to help people to do this is to move them out of a group where they have had quite an intense experience, either positive or negative, and dilute this a little bit by getting them into a different setting. Now I think if that had been made clearer to people, they might have made different decisions: (a) about what groups they went into, and (b) about whether Group B wanted to stay together.

(8) The NISW might find it very useful to offer the workshop to social work teachers, practitioners, and other social service personnel within a more limited geographical area. An intense examination and discussion in one area, involving a variety of agencies, could bring about a very strong impetus toward the delivery of services in a more unified manner. It could, certainly, further experimentation with, and evaluation of, more unitary approaches

to social work practice.

(9) Anything which the NISW can do to make reading material on unitary approaches available to intending workshop participants would probably heighten the initial level of understanding brought to future workshops. Further, if copies of relevant textbooks and papers are made available during the workshop, it is likely that they will be used.

CHAPTER VI

STUDY 1 AND THE CLASSICO-EXPERIMENTAL PARADIGM

In this Chapter, and the next, we will discuss applications of the paradigms in the foregoing Studies and further examine some of the paradigmic strengths and limitations indicated by such examinations. A major segment of the investigation reported in Chapter IV was intended to take the approach characterized as "classico-experimental", while that of Chapter V assumed the "socio-anthropological" perspective. In this Chapter we will enquire more closely into certain features of the classico-experimental paradigm which are exemplified by, or deficient in, Study I. In the next Chapter we will essay the same sort of exercise with regard to the socio-anthropological paradigm, using Study II as something of a specimen. In both Chapters, the paradigms will be considered in the light of experience during the Studies. Before we undertake these studies, however, a caution is in order. Even a cursory reading of the Studies should disabuse the reader of any notion that they are exemplary applications of their respective paradigms. Neither is put forward as an exemplar. Study II, perhaps, better typifies its paradigm because the socio-anthropological approach is not so rigidly delineated a methodological "package" as the classico-experimental, thus making it easier to manoeuver within its boundaries. Whereas the classico-experimental approach

can be presented as a relatively unambiguous set of prescripts, containing a number of fairly well-defined procedural steps, the socio-anthropological paradigm appears to be much more loosely-knit. As a relatively naive post-graduate student with only my own efforts to rely upon and very few resources at my disposal. I found it easier to apply the less rigidly conditioned socio-anthropological approach than to meet the more stringent procedural demands of the classico-experimental paradigm. This observation might, in fact, suggest a first tentative conclusion: that the socio-anthropological approach is somewhat more adaptable in application because its directives are not as inflexibly demanding of scarce resources; but more of that later. We will pay attention to some of the dimensions upon which the Studies can, and cannot, be validly seen as representative of the principles and prescriptions associated with their respective paradigms. If we can establish that, in at least a limited number of ways, the dimensions of the Studies accurately reflect paradigmic dimensions, then we might examine each in terms of the qualitative criteria outlined in Chapter II.

We noted that neither Study even approaches an "ideal" representation of its paradigm-in-action and suggested that one reason for deficiencies was my own personal and academic short-coming, coupled with a number of rather practical contextual limitations. Since both Studies resulted almost entirely from my efforts alone, they are bound to suffer from my own inadequacies; inexperience in

using the paradigms, gaps in substantive knowledge, unrecognized biases, and lack of resources. While I did have experience as an instructor and student prior to these investigations, my experience as a formal "course evaluator" was more limited. Preparatory reading and discussion with others, as well as some of my previous academic preparation and the knowledge gained from a pilot study, had provided me with a basic understanding of the tasks I was to undertake; but I was not an "expert" evaluator. In addition, financial, temporal and social limitations sometimes made it impossible, and more often impractical for me to more closely approximate the demands of the paradimgs. As this discussion proceeds I shall try to indicate areas in which limitations arising from my own particular situation gave rise to weaknesses in both studies. At the same time, I hope to indicate that the paradigms themselves can be seen to suffer from a number of limitations and inadequacies. It is upon paradigmic strengths and weaknesses that we most wish to concentrate if we are to engage in assessments of their respective values to us as evaluators of social work instruction. Finally, I would suggest that defects in the Studies which are attributable to me and my specific situation may, at times, <u>interact</u> with paradigmic failings to effect even more serious inadequacies. I shall try to point out some of these interactional linkages.

Study I as a Specimen

In Chapter III we outlined eight steps by which the "ideal" classico-experimental investigation would proceed: (1) definition of objectives, (2) designation of treatment and control groups, (3) development of instruments, (4) application of instruments to all subjects, (5) instructional "treatment", (6) re-application of instruments, (7) data analysis, and (8) conclusions. Although it was necessary to omit Step 2, thus restricting activity in Steps 7 and 8, part of Study I represents an attempt to follow these procedural prescriptions. Examination of the paradigm's application in this instance will therefore proceed in likewise fashion.

- Definition of objectives

In order to undertake a study using the classico-experimental approach it was first necessary to examine the goals toward which the workshop was directed and to develop a set of appropriate learning objectives which were defined as specifically as possible. To this end the workshop's tutors were consulted and the ensuing discussion suggested that three major workshop goals were intended. It was the tutors' stated intention that the instructional sequence should assist participants in: (1) developing a theoretical understanding of the principles and dimensions of a unitary approach to social work practice, (2) learning to utilize a unitary approach in the "assessment" phase of social work practice, and (3) learning to utilize a unitary

perspective in developing interventive strategies within social work practice. However, although these were seen as identified goals for the workshop, tutors did not intend to didactically "teach" toward them as such. They appeared to be primarily concerned with promoting the establishment of a learning environment within which participants would have an opportunity to move toward achievement of these goals by their own efforts. Relative satisfaction with previous workshops suggested that participants would "naturally" tend toward achievement of the goals if they became immersed in a context which encouraged them to inquire into, discuss, and symbolically manipulate elements of the unitary approach. It would seem, also, that tutors' goals were not as clearcut as they originally appeared to be. On the one hand, they possessed relatively well stated, generalized aims concerning the understanding of conceptual dimensions and issues, and the acquisition of practical skills using a unitary approach to "assessment" and "intervention". On the other hand, something more seemed to be expected--something which could be stated only in rather vague terms. This further set of aims appeared to involve a desire that students would make their own critical assessments of the unitary approach in its various forms. Perhaps this intention could not be stated so clearly because some elements of it were vaguely seen to be in partial contradiction to the three stated goals. The resolve that students should come to "understand" the unitary approach and learn to use it in particular ways appears to assume that there is

basic merit in the conceptual entity. On the other hand, if it is intended that students should explore the relevant concepts and reach their own conclusions with regard to merit, the initial assumption would appear to bias the exploratory process. This confusion was found to be a rather mystifying element in the structure of the workshop's printed program as well. The written program was sub-divided in such a way that it tended to emphasize the "understanding" and "learning to apply" tasks implied by the stated goals. At the same time, the amount of time delegated to small group discussion and consolidation in the plenary sessions would appear to encourage individual exploration and arrival at conclusions which might lead to a rejection of the unitary approach. Tutor intentions were, therefore, not entirely clear.

In order to use a classico-experimental approach to evaluation, learning objectives had to be specified as clearly as possible. At the same time, the somewhat confused nature of tutors' intentions suggested that one could not try to measure all of the major objectives and simultaneously maintain a substantial measure of logic. Arguing from Bloom's taxonomy of the "cognitive domain"--where evaluative behavior is represented as a high order class of cognitive behavior--the assumption was made that at least a part of the evaluative intention might be subsumed under the "understanding" objectives.[1] Thus, the three originally expressed goals were used as criteria against which the outcomes of the workshop would be assessed. It was intended that

assessments with regard to the "understanding" criterion might include portions of the second, more evaluative constellation of tutor goals. Since "understanding" was also assumed to be enabling of application, that objective took on a signal importance.

This paper has previously suggested, more than once, that instructional programs are characterized by a wide variety of objectives, arising from diverse sources. Furthermore, it would appear that student-held objectives are rather important determinants of events that actually occur during an instructional sequence. Study I necessarily ignored students' objectives and concentrated on tutors' objectives because to do otherwise would have been entirely impracticable. The classico-experimental paradigm does not, in fact, encourage examination and utilization of student-held objectives. As was the case in Study I, the final list of selected students is often not completed until relatively "late in the day". In Study I the list of students was not available until approximately one month prior to the workshop. If pre-test instruments were to be constructed and administered before the workshop began, there simply would not have been enough time available to base any measurements on objectives provided by the students. There would not have been even enough time to adequately enquire into the objectives held by twenty-eight individuals, spread all over Britain, who could not be gathered together for exploratory discussions. Tutors' objectives were available early enough in advance of the workshop to allow

for instrument development and pre-test application; students' objectives were not. Similar situations are very common, especially in short courses of continuing education or in-service training. Membership lists are often not finalized until only just before the instructional sequence begins. In the extreme, programs are sometimes run for "anyone who shows up", and training officers or continuing education instructors will testify that it is often impossible to predict exactly who will attend a course until the very day, sometimes the very hour, that it begins. In such curcumstances, it is virtually impossible to explore student learning objectives and develop measures of their achievement in time for a pre-test to be administered. It might be argued that the goals of applicants to a course, or a sample of possible participants, could be examined early enough to yield valuable information on student objectives. In certain instances, for example where a waiting list existed, this would appear to be a feasible alternative. It would, of course, entail a further expenditure of resources to poll and examine the objectives of prospective students, and it seems unlikely that such a sample drawn in advance of the student selection procedure, would accurately represent the objectives of students who were eventually selected to attend. However, in longer instructional sequences and instances in which prospective students display relatively common objectives, such an investigation might prove very useful. In such cases, the learning objectives arising from a thorough needs analysis

could serve as one set of criteria against which instructional effectiveness might be comprehensively assessed. Clearly, however, in a substantial number of instances the classico-experimental paradigm is not well suited to the use of student-held objectives. Almost out of necessity, it must concentrate on instructor-held objectives if pretests are to be available for appropriate administration.

To the extent that the tutors' major goals were used as criteria for judging the effectiveness of the workshop, Study I conforms to the ordinances of the classico-experimental approach. It should be noted, however, that learning objectives were not expressed in terms of discrete student behaviors which are observable by reference to objectively verifiable indicators. The level of required cognitive complexity implied by the goals made close specification a singularly difficult undertaking; though it may be argued that through a further expenditure of time and effort, objectives could have been beneficially partialized. In the event, the only instrument which appeared to offer a practical method of assessing goal achievement relied on the judgement of experts. Further partialization and specification was not seen to be practical.

The impracticality of objective specification in Study I should not be taken as a blanket condemnation of the approach. Rossi and Williams note that though it is difficult enough to change individuals, it is even more difficult to change them to an unspecified state.[2] Perhaps even more to the point, it is more difficult to <u>demonstrate</u>

that individuals have been changed to an unspecified state. Clearly, the operational definition of objectives encourages strict attention to the intentions that accompany instruction. It makes possible a concentration on assessing the extent to which stated intentions are met by the instructional program. Such assessments of effectiveness must be a major concern in any attempted program evaluation. Furthermore, attempts to gauge the extent of effectiveness by using a set of commonly-agreed-upon measurments which reflect as much objectivity as possible, appear to be laudable ventures. Validatory efforts are extremely important to meeting demands for accountability which arise both within and without the profession. Additionally, they may provide useful information for the instructor or administrator who must decide upon broadly-drawn action alternatives. If a particular practice or circumstance does not appear to be very effective in producing desired outcomes, a change in practices or circumstances is indicated. The operational definition of objectives, then, appears to underpin worthwhile assessment activity which may inform decision-making in such areas as program continuation/discontinuation, accountability, and broadly-conceived program improvement.

At the same time, it would appear that quite a number of the important intended effects of our instructional programs will be practically unmeasurable. This is not to say that such effects will be theoretically unmeasurable; theoretically, all effects are measurable. However, even if

it should be amenable to definition in terms of an available operational test, an effect may be entirely too difficult to measure given the resources available and within the environmental circumstances that exist. Many of the problems in Study I resulted from limitations imposed upon the investigative resources available. Much desirable measurement activity had to be considered impractical.

Facilitation of objective measurement, however, is not the only benefit to be had from efforts to operationally define objectives. As in the Unitary Approach Workshop instructional programs are characterized by a diversity of aims which arise from a variety of sources. It may be, for instance, that one instructor within a teaching team will have different instructional intentions from those of his colleagues. Students may have aims different from those of their instructors. A departmental staff training authority may hold instructional aims different from those envisaged by departmental directors. Attempts to specify objectives in operational terms may do much to uncover these differences and facilitate less ambiguous communication about curricular intentions. Further, attempts to operationally define instructional objectives may provide a framework within which the negotiation of mutually desired goals can be further explored.

It is likely that a more detailed objective specification will have other benefits in the curriculum planning process. While clear communication and negotiation with regard to instructional intentions should help to promote

participant satisfaction with the curriculum plan which is ultimately adopted, efforts to operationally define objectives are also likely to provide explicit criteria for selecting, rejecting or modifying instructional content and method. Program structure may be more closely related to instructional intentions. For example, a curriculum plan to guide an instructional sequence concerned with social group work modes of intervention may specify that students should come to understand theoretical constructs such as "social norm", "shared group values", "functional roles", "boundary maintenance", and "subgroup formation". A comprehensive set of instructional objectives will specify these content areas and, perhaps, suggest means by which the learning of such content may be facilitated.

Lest the circumstances surrounding Study I lead us to consider problems encountered there as being entirely unique, it should be recognized that difficulties will arise in any approach to evaluation which begins from the specification of instructional objectives. Any pre-specification of outcomes which serves as a focus for investigative activity may introduce an over - narrow investigative focus; other equally significant effects may be obscured, overlooked or ignored. Thus, significant but unanticipated or atypical effects may be left unattended. To the extent that such other outcomes represent crucial achievements or harmful "side-effects", an investigative approach which does not conscientiously seek them out is likely to leave the decision-maker ignorant of valuable information. As Scriven

notes, goals represent only one subset of anticipated effects and an evaluation with respect only to goals does not consider all of the anticipated effects, much less those that are not anticipated. Since such an evaluation affords a very limited profile of the program under study, Scriven is led to propose evaluation which is "goal-free".[3]

An associated area of difficulty is the tendency to assess effectiveness in terms of the outcomes which are conveniently measured. For instance, it is often easier to assess the learner's retention of specific "facts" than to estimate his understanding of a complexly interrelated set of concepts. It may be much easier to assess his skill in applying knowledge to the analysis of a hypothetical case study than to assess his skill in effectively intervening in a real-life problematic situation. Efforts to operationalize objectives may lead to investigative results which place disproportionate emphasis on easily operationalized objectives. In many cases, though we may be able to develop precise measures of immediate objectives, it is the more long-range objectives--which cannot be practically operationalized--that represent the most significant or desirable outcomes. As Bennett and Lumsdaine observe: "This can place us in the strange position of having to weigh indices of program impact in inverse proportion to the reliability with which they can be measured."[4]

- Control Groups

An aspect of the classico-experimental paradigm which is not well represented by Study I is the use of a control sample. If properly constructed and applied, designs incorporating control groups can deliver assessments of effectiveness containing evidence which permits substantial inferential capacity. In an investigation like Study I where little change between pre-test and post-test was indicated, the lack of a control group may not appear to be such a serious shortcoming. In such a study, the use of a control group design will probably provide negligible further amounts of valuable information. However, had we found great differences between pre-test and post-test, control group information might have given us a relatively strong basis for discerning the extent to which participation in the instructional program contributed to those changes. At the same time, it should be noted that information is required not only on control sample performance and "baseline" characteristics, but also on control group activity during the instructional period and contextual influences. If, as in the CARE Project study cited earlier, we gather information only on pre-test-post-test changes, our inferential capability will remain rather low. For, we cannot assume that control subjects did absolutely nothing while the treatment group participated in the instructional sequence. In cases where the control group is participating in another form of instruction, the investigator should be aware that, in reality, he is comparing the putative

effects of one sequence of instruction with those of another; information about both programs is vital. In cases where control subjects are not engaged in an alternative instructional sequence, it cannot be assumed that they were not engaged in an activity which affected their performance test results. Control subjects may have undertaken self-initiated individual study; they may have undergone work experiences which affected their post-test performance; they may have inadvertently acquired relevant skills, knowledge, or attitudes in a host of other ways. Obviously, the investigation which does not enquire into such activities renders little useful information from the inclusion of a control sample.

There are definite inferential benefits to be had from the use of a control sample. For example, even if an investigator does find scarce evidence of performance differences over time, more detailed analysis may indicate that the instructional sequence had very significant effects for certain subgroups of subjects and not for others. Even if overall results comparing treatment and control subjects display no marked differences, the finding that a particular subgroup of treatment subjects performed significantly differently from a comparable subgroup of "controls" may be a valuable indication of differential effectiveness.

While it is obvious that the design of Study I did not include a control group, and it is possible that a control sample would have extended our inferential capacity, it is also felt that little apology for this deficiency is

necessary. Under the circumstances, use of a control sample appeared to be out of the question. There existed no large "pool" of applicants from which an appropriate control sample might have been drawn. Furthermore, there did not appear to be any basis for assuming that workshop participants comprised an accurately representative cross-section of British social workers in general, thus leaving no basis for assuming that a randomly drawn sample of British social workers would suffice as a control group. This is not to suggest that a matched sample control group could not have been composed once the characteristics of attendees were investigated. Obviously, each workshop participant might have been matched with a similar subject from the universe of British social workers; with regard to any number of personal, social, and professional characteristics which one might have considered important. However, locating and enlisting the cooperation of those control group subjects, as well as applying instruments to them, would have been entirely impractical under the circumstances. The resources required to do so were simply not available. Further, even with almost unlimited resources, it would have been very difficult to find appropriate control group subjects in sufficient time to allow for pre-test application at the same point in time as for workshop participants. Assessment at some other time would have introduced a number of other uncontrolled variables and, therefore, would not have resulted in a large increase in inferential potentiality. This problem appeared to be largely due to

the current status of the unitary approach in British social work circles. Acquaintance with the elements of a unitary approach to social work practice amongst British social workers, particularly social work instructors, had been developing very rapidly at the time of the workshop and might well have been a significant uncontrolled variable within any "control" group selected. Further, even if a great deal of money and effort had been expended on finding and using an appropriately matched control sample, the "motivation" variable could not have been stringently controlled. Workshop participants were, for whatever complex of reasons, so motivated to attend that they went to the trouble of applying, as well as arranging personal, social, and working circumstances to enable attendance. It is extremely doubtful that the same level of motivation could have been "built in" to any control sample unless a large pool of applicants who were prepared to attend was available. And if, as suggested in the literature, motivation was to be measured by action rather than by intention, only one indicator of motivation, the act of application, would have been available in any event.[5]

In discussions on curriculum development and evaluation, the question of student motivation is very often ignored. Growing concern with this deficiency, exemplified by use of the concept "praxis" rather than "process", should give us reason to reconsider motivation as a potent influence on human behavior--including learning behavior.[6] As Taba says, "Intrinsic motivation--curiosity, sensing the relevance

and purpose of what is being learned, a sheer drive to understand, or a quickened awareness--is likely to be a more stable stimulator of learning than extrinsic rewards. Learning engendered by intrinsic motivation is more likely to be retained and used again."[7]

This concern with the differential influence of personal motivation levels suggests only one major dimension of a broader set of difficulties; difficulties which seriously circumscribe the real value of research designs incorporating "control" samples. The problem set might be labelled "self-selection difficulties". As Weiss notes: "People who choose to enter a program are likely to be different from those who do not, and the prior differences (in interest, aspiration, values, initiative) make post-program comparisons between 'served' and 'unserved' groups risky."[8] It is very often the "best risk" individuals who select themselves into a program; not those who <u>most need</u> the program's services. Thus, when program participants display better post-program outcomes than control subjects, it is entirely possible that such individuals would have done better in any event--even without any program influences. Through the use of "unawares" (individuals with similar characteristics who were not aware of the program) or "geographic ineligibles" (individuals with similar characteristics who reside in areas with no access to the program), it is sometimes possible to construct a relatively valid control sample. However, this may easily become an expensive activity--definitely too expensive in Study I, for example--

and a number of difficulties will remain. One might ask, for example, what it was about the "unawares" that blocked their knowledge of the program. One might question the degree of similarity of environmental conditions in a location which did not offer access to the program. Clearly, complex questions arise in any attempted solution to this set of problems, and with each set of solutions will come increases--often substantial increases--in research costs.

The major point was that inclusion of a control group in the design of Study I, while extending the possibility for greater inferential strength, was not seen to be a feasible proposition under the circumstances. In order to locate appropriate subjects, enlist their cooperation, and apply instruments in the prescribed manner, one would have required much more time, money and energy than was available.

The circumstances surrounding Study I do not represent a unique example of a situation within which use of a control sample is impractical or of questionable value. Very often, no population pool of appropriate subjects is available for composing such a group. In instructional programs where the number of applicants far exceeds the number who can be admitted to the program, construction of a control sample may be possible; but, even then, great care must be taken. The very fact that one group of students was selected to take part in an instructional sequence while others were rejected may indicate that significant differences between the two groups already exist. In graduate social work programs, for example, if selection procedures are effective,

we might expect that those students selected to participate in the program are more capable of benefiting from the instruction, more sympathetic to the values espoused by professional social workers, and able to score higher on intelligence tests.[9] If "treatment" group subjects do have a significantly higher potential for learning, or for scoring higher on the instruments used in assessment, comparison with lower potential "control" group subjects may be quite misleading. Clearly, student selection factors may operate to render a control group design invalid. The investigator who does not attend to such factors may end up comparing apples to oranges.

Such concerns are a major reason for randomization in treatment/control group composition. If students can be assigned to groups randomly, significant intervening variables like selection criteria may be "controlled for". However, random assignment often heightens other concerns; particularly ethical concerns.[10] Ethical reservations about exposing a group of students to programs which are expected to be ineffectual, possibly even harmful, are understandable.[11] In the extreme, one cannot admit students to an instructional program and then provide no instruction because they have been designated "controls". Similarly, time or "unlearning" requirements may make it impossible to rectify losses incurred by uninstructed control subjects. Although it may be argued that ". . . ethical reservations about subjects getting the 'wrong' treatment can be met when unsupported by objective evidence about the superiority of

one type of treatment over another", prospective control group subjects or instructors entrusted with assisting them to learn may not be so easily convinced.[12] In most cases, evaluation activity would not have been initiated were it not suspected that one instructional sequence was somehow a "better" alternative. Furthermore, instructors are often opposed to random assignment of subjects. They wish to select students by making professional judgements about a variety of suitability factors, not by chance. Finally, concern is generated when students are assigned to treatment or control groups without their consent. Although methodological contingencies sometimes require such an approach (e.g., to control for "Hawthorne" effect), its ethical integrity may be dubious. As Clarke and Cornish note: "It often seems that least effort is made to gain the consent of those who are likely to have least reason for wanting to give it."[13]

The various additional expenses associated with control sample composition and use in Study I rendered such a design impractical. Any control group design will require that the expenses involved in instrument application and data analysis will increase, but great additional expense may also be incurred in locating control group members, matching, and assessing control subject experience during the treatment period. Thus, though there may be significant inferential benefits to be gained from use of a randomized or matched control sample, these additional expenses, ethical difficulties, or logistical problems may leave such

a course of action impractical or undesirable. Furthermore, the inappropriate use of randomized controls may result in grossly misleading research findings. Such contraindications have led Edwards and Guttentag to conclude:

> Some control groups do not control. Random assignment of subjects to conditions often turns out not to have been random. We strongly agree with the argument . . . that randomization is best where possible. But we must add, 'It's nice work if you can get it'.14

- Development of Instruments

The two case study "exercises" in Study I constitute pre-test and post-test components of a performance test and represent an attempt to assess the extent to which the tutors' three learning objectives were achieved by students. The level of generality with which these objectives had been stated, combined with the high cognitive complexity implied (especially the "understanding" objective), had necessitated a search for some means of assessment that would encompass a large proportion of what the tutors actually meant. Thus, the tutor rating procedure was chosen. A method more consonant with the principle of operationalization in "objective" terms would have been to break down tutors' goals into a larger number of more "elemental" indicators which, when combined, might have provided more empirically-based evidence with regard to goal achievement. Critics would argue that my unwillingness to specify objectives in more distinctly behavioral "bits" constituted a significant departure from the prescripts of the classico-experimental paradigm. In ideal circumstances this might be considered

a strong indictment against the Study itself. The fact is, however, that it was not simply laxity which gave rise to the sorts of objectives and instruments used; again, practical considerations militated against further specification. Much closer scrutinizing of the objectives and differentiating them to accord with the development of more empirically observable indicators would have required an inordinate expenditure of time and effort, both on my part and on behalf of workshop tutors. Furthermore, the use of tutors as expert judges appeared to offer the investigation a degree of validity which it would have lacked otherwise. The tutors' ratings of case study "notes" could be expected to reflect the dimensions of their learning objectives much more adequately than could any necessarily circumscribed list of highly specified indicators. Specification of a comprehensive set of indicators for goals such as these was out of the question and, even under the best of circumstances, would not have provided a more valid reflection of tutors' intentions. Thus, the expert judgements of workshop tutors were viewed as representing highly valid indicators of the extent to which students demonstrated achievement of tutor-held objectives, and a practical approach to assessment. To the extent that this procedure attempted to assess the intended outcomes of the workshop by measuring objective achievement through performances tests, it fits squarely within the classico-experimental tradition. The choice of tactics may therefore be seen as both a response to practical limitations and a way of resolving some tensions between

validity and reliability requirements.

The use of two student rating instruments ("Daily" and "Overall" Evaluations) arose in Study I because the performance test procedure was not considered to be entirely adequate in producing required information; earlier we noted that similar designs are relatively common. In this case, it was felt that complete reliance upon tutors' judgements would provide insufficient information for the grounding of improvement decisions. Further, it would produce absolutely no information from the students' point of view and, although student rating scales represented a departure from use of the classico-experimental approach, rigid adherance to that paradigm appeared to render the investigation entirely too limited. This point will be taken up in later discussion.

- Application of Instruments (Pre-Test)

The case study "exercises" were initially applied to subjects after randomly selecting one half of the group to receive Case A and the other half Case B. This was done so that any differential effects on outcomes produced by the contents of the case studies themselves would be apparent. Prior to the workshop each prospective participant received one of the case studies, a set of instructions pertaining to the "exercise", and an explanatory letter. Each was requested to analyze the case situation and produce a set of "notes" which he would use in a hypothetical student-instructor seminar. The "notes" were to be delivered to me

upon arrival at the workshop centre. In the event, this procedure worked extremely well. Of the 28 workshop participants, 26 arrived bearing a set of "notes" to be handed in. Had return by mail been requested, it is unlikely that such a high proportion of returns would have appeared.

At the same time, the procedure was somewhat lacking in uniformity of application. Certainly, all subjects received case studies and instructions in the same way, but no effort was made to enforce the condition that only one half hour should be expended on task completion. It is entirely possible that one subject might have completed the task in ten minutes while another might have studied, ruminated, and laboured over it for two weeks. Furthermore, the student list provided by the NISW, though usually including office addresses, occasionally only provided a subject's home address. It is quite conceivable that subjects who received the "exercise" at home tended to treat it differently than those who received it at work. There were probably other relatively significant variations in application as well. The point is that the circumstances under which instruments were applied did vary somewhat from subject to subject.

The instruction sheet which accompanied case studies proposed a hypothetical situation with which, it was hoped, all participants would be able to identify. Since all workshop members were experienced social workers, it was assumed that the analysis of written case material would not be an entirely new task to them. One might, however, argue that

subjects who led student seminars in the course of their normal employment would be better able to identify with and complete the task than those subjects who had never engaged in such activity. Varying the forms of the task for varying types of subjects was considered but eventually rejected in favour of uniformity. Further, it was hoped that the "one half hour" stipulation would make the "exercise" appear to be easy enough of accomplishment that no subject would be initially deterred from undertaking it, and that this limiting condition would impose some control on variability amongst subjects with regard to the amount of effort expended.

Finally, subjects were informed that only the workshop evaluator would know who produced any particular set of "notes", thus indicating that a high degree of confidentiality would be observed. Although little empirical evidence has been produced to support the assumption, it is generally believed that assurances of relative anonymity will result in a higher proportion of returns. Since there was no need for the identification of particular individuals by anyone other than myself--one strength of the expert judging procedure, in fact, rested on subject anonymity--this assurance was not difficult to uphold.

In summary, although the conditions under which instruments were applied were far from laboratory-like, efforts were made to standardize them to the extent possible. The use of a mail-out instrument dictated against stringent control over the circumstances of application but no especially significant variations were apparent and more

rigid control would have required either massive expenditures of resources or interference in the instructional program itself. Under this set of circumstances, the requirements of experimental control in instrument application prescribed by the classico-experimental paradigm were practically observed.

- Instructional Treatment

The classico-experimental paradigm appears to concentrate almost no attention on the actual implementation of the instructional program. We have already noted that this is one of the areas in which the paradigm has been subjected to criticism. Since the approach does not prescribe collection of details about dimensions of the instructional sequence itself, it was decided that Study I should not dwell on this area. At the same time, it was felt that some exposition of the "treatment" variables was necessary if the evaluation report was to provide any information whatsoever on interventive activity. The paradigm's apparent neglect of this dimension will necessarily concern us in later discussions. For now, it should simply be noted that very little is expected of the classico-experimental evaluator with regard to "treatment" descriptions.

- Re-Application of Instruments (Post-Tests)

The second component of the performance test was applied approximately two weeks after the workshop was completed. Subjects who received Case A for the pre-test, received Case B, and vice versa. This time, however, reply by post was

necessary--introducing another possible source of bias into the data. Inclusion of a stamped and pre-addressed return envelope with each set of instruments, as well as the use of reminder letters for those who did not immediately return their "notes", probably helped to minimize biases arising from resistances to a mail-back return. Eventually 18 sets of "notes" were gathered. Unfortunately two of these were from subjects who had not responded to the pre-test, so only 16 complete pairs of "notes" were available for analysis and judging. This response, over 60 percent, is congruent with proportions of subjects usually expected to respond to mail-back research instruments. In fact, it is a good deal higher than the response achieved in many studies; a fact which may be accounted for by the relatively high level of personal contact possible in a small-scale study of this sort.

The same controls over the conditions of application which we noted in discussion of the pre-test applied on the post-test. The same instructions were given and, other than the postal reply requirement, the same possible sources of variability were present. It might, however, be hypothesized that since subjects were unlikely to have face-to-face contact with me again, they felt less threat in not responding. During the pre-test they expected that we would meet and that I would expect them to hand over a set of written responses. During the post-test this threat was not present and its absence may be regarded as a possible source of bias which did not exist in the pre-test situation.

In any event, the application of post-test instruments completed Step 6 of the classico-experimental paradigm's ordinances.

Before we leave this area of discussion, however, some further comment on "before-and-after" designs seems appropriate. Observing indicators of criteria achievement prior to and after an instructional sequence, of course, makes it possible to gauge the extent of change which occurred during the instructional period. It is folly to assert that change has taken place if assessments are conducted only at the end of instruction. On the other hand, Study I suggested at least two major associated difficulties. In the first place, it would appear that practical difficulties may be encountered in employing any but the most readily available participant intentions to form instructional objectives. This was discussed earlier when we noted that student-held objectives had to be ignored in Study I because the opportunity to inquire into them was not available. In many instructional programs, knowledge of student-held objectives may require an examination of needs and desires which simply cannot be accomplished prior to program onset. This would appear to be a particular problem in the case of in-service training or continuing education programs where the time span between student selection and instructional onset is often extremely short and where adequate needs/desires analyses have not been carried out as part of the curriculum design process. The requirement that all assessment methods to be used must be available for pre-test

application may effectively limit the use of certain, possibly critical, assessment measures.

The second area of difficulty arises from the immutibility of the before-and-after design once initiated. As an instructional program operates, it may be expected that unanticipated changes will occur. The objectives and values of program participants may change in response to program experience and changes in external factors. The program, itself, may change shape and character as it is instituted. External circumstances, to which the program was a response, may change. Knowledge of program consequences accumulates. A paradigm for evaluation which is not adaptable to continuous sympathetic change may not provide adequate information to decision-makers and may, in fact, incur serious resistances in program participants. This is not an obvious difficulty in Study I because we were dealing with a short-term instructional program in which change was not so apparent. However, in a more long-term program, difficulties might be much more obvious. Actually there are two sources of problem here. The first, already alluded to, is the fact that we are usually not so interested in evaluating the theoretical design of the program, rather in evaluating it as implemented. Second, instructional changes that become obviously necessary or commendatory during the course of a program often cannot be inhibited for the sake of the research effort. If the research design requires permanence in a situation where change would otherwise be warranted, it is likely to encounter

ethical difficulties and engender disenchantment amongst program participants.

-Data Analysis

Before the sixteen pairs of "notes" generated in Study I's performance test could be submitted to the panel of experts (the tutors) for judging, it was necessary to transcribe them into a relatively standardized form which would minimize the possibility that individuals might be identified or that differentiation between pre-test and post-test responses might be possible. All thirty-two sets of "notes" were therefore rewritten in standardized form and labelled only by randomly chosen identification numbers. Great care was taken to ensure that the standardization process affected only the form in which "notes" were presented, not the content contained by them. They were then sent to the panel of four tutors who were instructed to rate each set on three dimensions, each dimension employing a scale which ranged from 1 to 8. A rating of 1 or 2 was designated as "very poor", while a rating of 7 or 8 was said to represent "very good or excellent". Rating tables were supplied and the three dimensions outlined as below.

(1) "Writer's demonstrated understanding of the unitary perspective."

(2) "Writer's demonstrated ability to apply a unitary perspective to analysis and assessment of the case material."

(3) "Writer's demonstrated ability to apply a

unitary perspective to planning of interventions in the case situation."

Judges were provided with copies of the case studies and instruction sheet sent to subjects, and each was asked to rate 18 sets of "notes". This meant that each judge would rate both sets of responses from nine subjects, allowing two sets of rated pairs for 14 subjects and four sets for two subjects. Although it would have been desirable for all 16 pairs to be rated by each judge, the amount of time and effort involved for the tutors rendered that large a task prohibitive.

Upon completion of judging, rating data were compiled and analyzed. Differences between pre-test and post-test scores were subjected to t-tests to determine the statistical significance of observed changes, and a large number of correlations were computed. Performance test data were also correlated with data obtained from the student rating scales and participant information sheets. Any results which appeared to be statistically significant--as well as a number which did not--were rechecked at least once and recorded centrally. Thus, Step 7 was completed.

-Drawing Conclusions

The final Step specified in our outline of the classico-experimental paradigm was to draw conclusions from the analyzed data. Although the lack of an appropriate control group initially appears to diminish the strength of conclusions about workshop effectiveness (had data from an appropriate control group been available, it might have been used to

strengthen any indications that the instructional sequence actually produced or inhibited intended changes in participants), the data which was available indicated that very little intended change had taken place between pre-test and post-test. Additionally, much of the apparent change which did register was not in the intended direction. It was only the availability of student rating data which permitted many of the conclusions drawn. If only the performance test data had been available, very few conclusions would have been possible and their tentative nature would have been even more apparent. Be that as it may, had the inclusion of a control group been possible, it seems highly unlikely that it would have substantially altered or supplemented the conclusions which were reached.

It would appear that the design and implementation of Study I do follow the set of prescripts which characterize a classico-experimental approach to evaluation research. Except for the necessary omission of Step 2, all Steps were taken in the prescribed order and experimental principles were upheld to the extent that circumstances allowed. It is suggested, therefore, that, though far from an exemplar, Study I represents a specimen investigation which roughly reflects many of the qualities characterizing the classico-experimental approach. The added dimension of student ratings is not a requirement of the paradigm and will, therefore, not be examined as an attribute of the classico-experimental approach. It will receive closer attention as a separate entity, at the end of this Chapter.

Quality Issues

In Chapter II we outlined four characteristics against which the quality of data and data-producing instruments may be assessed: (1) validity, (2) reliability, (3) generalizability, and (4) utility. It is intended that, in this Chapter and the next, these concepts will be used in examining the quality of the information and information-gathering media associated with the specimen Studies. Hopefully this will lead us to a more complete understanding of some of the strengths and weaknesses of each Study, and indicate further some of the advantages and limitations of the paradigms themselves.

Validity

Our considerations with regard to "quality" will necessarily cause us to pose a number of questions about the performance test used in Study I. It, after all, represents a major element reflecting the classico-experimental approach, and examination of its dimensions will direct us toward other important considerations. We might, therefore, begin by asking, "To what extent do the objectives around which the test was constructed validly represent the aims or goals intended." It has already been suggested that by stating the objectives at a relatively high level of abstraction we assured that they remained closely related to the goals put forward by the tutors. The tutors declared an intention that the instructional sequence should assist participants in understanding the dimensions and principles of a unitary approach to social work practice; further,

that it should assist them in developing abilities to assess and intervene in problematic social situations according to the dictates of a unitary perspective. So, if tutors' stated goals are considered to represent the only valid criteria against which course outcomes could be judged, the performance test procedure would initially appear to be highly valid. To the extent that the objectives say what the tutors meant, they are valid. As we noted earlier, however, it is most unlikely that the objectives stated encompassed all of the major goals held by the tutors. Apparently there were also tutor-held aims which had to do with providing an opportunity for participants to explore and evaluate dimensions of the unitary approach. To the extent that such an aim cannot be subsumed under the three stated objectives, the information gathered lacks validity. Since tutors did not appear to be entirely clear about all of their aims, validity cannot be comprehensively assessed.

It is also worth noting that there were, almost undoubtedly, tutor-held workshop goals which were not directly related to student learning and which were, therefore, not assessed by the performance tests. It is likely that tutors held a variety of diverse aims for the workshop (e.g., extending the NISW's sphere of influence, creating more widespread acceptance of the unitary approach, meeting the expectations of their employers and colleagues) that were only remotely, if at all related to their stated workshop objectives. Furthermore, we may be certain that individuals and groups other than the tutors (e.g., students,

NISW officials, participant's employers) had workshop-related goals which were not represented by the stated objectives. It begins to appear as if the stated objectives were far from comprehensive enough to be judged entirely valid representations of major workshop goals.

Proponents of the classico-experimental paradigm would argue that, to the extent these other aims and goals were important, they should have been specified and operationalized; and movement toward achievement of them should have been assessed. In the case of student-held goals, we have already noted that this is impractical advice. Limitations on resources make the specific definition of only a few objectives practical. It may be hoped, within the practical limits of the investigation, that indicators of achievement for the most important objectives will be included; but where there are numerous important objectives valid use of the classico-experimental approach rapidly becomes impractical. If the prescripts of the paradigm are to be followed, objectives must be stated and indicators measured in ways which usually require relatively large expenditures of resources; thus the statement of large numbers of objectives and the measurement of large numbers of indicators becomes a practical impossibility.

A second question which must be asked in order to assess validity is, "To what extent do the behaviors measured validly represent the defined objectives?" By asking the tutors themselves to make assessments, using their own definitions, a degree of validity is assured. The tutors

judged performance test responses according to their own conceptualizations of "understanding", "ability to assess", and "ability to intervene", and it is in these definitional terms that objectives were stated. Assuming that tutors' definitions remained relatively consistent over the time between statement of objectives and judgement of performance test responses, the rating data should retain a high degree of validity.

A further element of untidiness in the stated objectives suggests another validity problem. Although it was left unstated, it is likely that tutors intended for students to manifest their newly found abilities on the job. There would appear to be little value in attaining a practice skill unless it is in using that skill in "real life" situations which require it. The performance tests referred entirely to two hypothetical situations ("case studies"). If his response to a hypothetical set of circumstances indicates that a student has achieved an "understanding" of the theoretical principles of the unitary approach to social work practice, well and good, but if his response indicates an ability to "assess" or "intervene" appropriately in a hypothetical situation, does it indicate that he could appropriately assess or intervene in a real situation which presents itself in the course of his work? In the first place, the hypothetical case studies represent only one very limited segment of social work practice. Since the situations portrayed in the case studies were highly reflective of circumstances, events, and individuals

in actual social work cases, and since they appear to represent a certain sort of case situation which will commonly arise in the practices of many social workers, perhaps they are a relatively accurate reflection of certain social work "realities". To the extent that they are, the student who indicates an ability to react appropriately to them is likely able to react appropriately in a similar situation arising on the job. Obviously, however, these hypothetical case situations are quite dissimilar to uncounted problem-complexes presented to the social worker in the course of his practice; problem-complexes within which the tutors hoped he would be able to apply a unitary approach. The student's ability to respond appropriately to the dissimilar hypothetical case study would not appear to provide much indication of his ability to respond appropriately under these other conditions. For example, the ability to respond appropriately to the case studies used in Study I may not indicate an ability to respond appropriately in a situation where two large community groups are involved in a competition for power which is destroying neighbourhood peace and solidarity.

In like manner the task set in the performance tests may not accurately represent the tasks implied in the objectives. The performance test in Study I casts each subject in the role of a fieldwork teacher making notes on a case study so as to better lead a seminar discussion with his students. However, the objectives imply an ability to use a unitary approach in actually dealing with problematic

social situations. Although the hypothetical seminar situation may be viewed as having problematic and social characteristics, it would not even fall within some definitions of "social work practice". The task was set with the expectation that a majority of workshop participants would be social work instructors, and the assumption was made that for such individuals, such a task would be part of their social work practice. Obviously the task is less appropriate to other sorts of social work practitioners who, hopefully, will use the unitary approach in other ways. This problem arose out of the heterogeneity of participants in the workshop and the intention that all subjects should be assessed with the same instrument. The "application" objectives were framed in broad terms but assessments could only be made on specific individuals.

We have seen that the behavior ostensibly measured by the performance test is not entirely representative of the behavior implied by the stated learning objectives. It is suggested that this is not only a weakness in Study I, but also an indication of weakness in the classico-experimental design. In any but the most simple of programs, not only would multitudinous objectives have to be stated to encompass the major aims and goals held by various parties, but a very large number of behavioral indicators would have to be used to girdle the behaviors implied by each objective. Unless very extensive definition of objectives and behavior measurement is carried out, therefore, the validity of the study is questionable. If high validity could be attained

it would necessarily be at very great expense.

Reliability

A different line of questioning asks, "To what extent has the procedure measured behavior accurately and consistently." The rather low levels of interjudge reliability found would suggest that the performance test used in Study I did not produce very consistent measurements. Put in another way, judges were rather inconsistent with each other in their ratings of performance test responses. This is amply illustrated by the ratings given in the two cases where all four judges rated the same material. It will be remembered that the "notes" from two (randomly chosen) subjects were submitted to all four judges. Table 6-1 shows the scores obtained on the "understanding" dimension.

TABLE 6-1

Four Judges' "Understanding" Ratings on Two Response Pairs

	Ratings			
	Judge A	Judge B	Judge C	Judge D
Subject A				
pre-test	6	7	4	4
post-test	3	7	6	1
difference	-3	0	2	-3
Subject B				
pre-test	7	6	6	5
post-test	6	6	3	2
difference	-1	0	-3	-3

While one judge gave subject A's post-test response a rating of 1 ("very poor"), another judge rated it at 7 ("very good"). While two judges' ratings indicate a considerable loss in understanding between pre-test and post-test (-3), another's indicates that there was no change at all (0), and the fourth's indicates an increase (2). In the case of subject B, two judges gave his post-test response a rating of 6, while another gave it only a rating of 2. His "difference" score varies from -3 to 0. Obviously judges did not demonstrate a great deal of agreement over what constituted an indication of "understanding" and what did not. Ratings on the other two dimensions were almost as inconsistent. The somewhat greater consistency on pre-test responses might suggest to some that the inconsistency displayed with regard to post-test responses was largely due to elements inherent in the particular case study used for the post-test. This is immediately discounted by the fact that subject A received case study A for the pre-test and B for the post-test, while subject B received study B for the pre-test and A for the post-test.

On the question of accuracy in measurement, we have no "objective" measures against which to compare the judges' ratings. Their very inconsistency is, however, an indication of inaccuracy. Although it is possible to be consistently accurate or consistently inaccurate, one cannot see how in this case it might be possible to be inconsistently accurate. For, even if one judge is "accurate", the other judges do not consistently agree with his ratings and,

therefore, must be "inaccurate". It must not, however, be assumed that consistency would indicate accuracy. Even if the data had shown very high levels of interjudge reliability, it would be entirely possible that judges were merely consistently inaccurate.

The only available standard against which we might measure the judges' accuracy in Study I is provided by the student rating data. On both the "Daily Evaluation" and "Overall Evaluation" instruments students almost invariably rated the amount they had learned in the top half of each scale--indicating a widespread and consistently expressed belief that they had learned a considerable amount from participation in the workshop. Of course this may not be thought to reflect a belief that they had attained any of the knowledge or skills represented by the tutors' three stated objectives, but the relatively high ratings given on Items 8a and 8d ("understanding" and "assessment and intervention") would suggest that these dimensions conform to the overall trend. The judges' ratings of performance test responses do not correlate positively with this indication of student learning. Neither were statistically significant differences found between pre-test and post-test responses. If we are justified in assuming that the relatively high student ratings on "learning" items are accurate reflections of the actual amounts learned, we are forced to conclude that performance test data were not at all accurate. The extent to which that assumption can be justified will be examined later in this Chapter.

If student rating data are not used as criteria for judging the reliability of performance test data, it would appear that in Study I and others like it, the term "accuracy" becomes virtually meaningless--at least having little or no meaning beyond that attributed by any particular individual. For how can information be considered accurate or inaccurate if there is no referent against which its accuracy may be assessed? In Study I, no "empirically ascertainable" referent to "understanding" was put forward, only the tutors' subjective impressions; and they were inconsistent with each other. Theoretically, this is not the fault of the investigative approach. Through the definition of objectives in empircally measurable terms, the classico-experimental paradigm intends to impart wider meaning to the concept of "accuracy". It attempts to provide a widely recognized, standard form of measurement. Thus, the paradigm specifies that measurements should be taken in particular ways and that the accuracy of measurements may be assessed by reference to the extent of compliance with prespecified techniques. If objectives can be operationally defined and instruments applied so that repeated measurements provide almost identical values, the procedure is said to be highly "accurate".

Thus, the classico-experimental paradigm does appear to offer possibilities for the production of reliable information. Earlier discussion, however, suggested that in order to obtain that information it may be necessary to invest large amounts of time, money and/or energy. Furthermore, we certainly will not be able to automatically assume

that because the data are reliable, they are in any way valid. In fact, in practical terms, the operational definition of objectives called for by the paradigm would appear to operate against the validity of data by decreasing its comprehensiveness in relation to the dimensions of stated goals. We have noted that in order for a set of specific objectives to comprehensively cover the meaning of abstractly stated goals like "understanding", it will, of necessity, contain a large number of specific objectives. We have further noted that comprehensive measurements with regard to the achievement of <u>each</u> objective may well require measurement of a large number of behavioral indicators. It was concluded that these necessities often render the paradigm impractical in use. It would, therefore, seem that there is a tension between the qualities of "reliability" and "validity", at least in the human sciences where measurement techniques are not well developed and often expensive in application. To the extent that a procedure strives for validity, it is likely to suffer in reliability; to the extent that it can become more reliable, it is likely to lose validity. In order to achieve high levels of both qualities, analogous levels of resource expenditure are necessary. The classico-experimental paradigm appears to concentrate more heavily on producing reliable data, often to the detriment of either the data's validity or the paradigm's practicality in operation.

Complementary with a concentration on reliability, the classico-experimental approach favours the generation

of quantified information. This emphasis on measurement--assigning numericals according to rules--tends to reinforce the reliability of the information. Quantification encourages uniformity in observation as well as clarity and economy in communication about observations. Through operationalization, the measuring procedures employed become incorporated into the very definitions of concepts, thus providing empirically observable referents to meaning. The use of such measures, then, assists greatly in cutting through contradiction and confusion; especially in answering questions like "how much" and "how many". As a result of using such measures, mankind has benefited greatly. As Ackerman points out, they have ". . . facilitated the transmission of data and technique, promoted an incalculable contribution to the sum of knowledge, and brought about an extraordinary refinement of our critical faculties."[15]

Study I illustrates one strength of quantification very well. The quantitative nature of the data made it possible to compute statistical coefficients of interjudge reliability which indicated that the information was not particularly reliable. It also made possible statistical correlations with student rating data and use of the t-test to determine the statistical significance of changes supposedly induced by workshop participation. Thus, with quantitative information, analysis of the data may be more precise and it may become relatively easy to assess the information's reliability.

There are those who would suggest that the weaknesses imposed by quantification sometimes outweigh the strengths. The easy assessment of reliability by statistical methods is not of genuine benefit if the quantified nature of the information produced unreliability or invalidity in the first place. Further, quantification which permits statistical manipulation and ease of data examination, does not necessarily produce more meaningful data. It may, in fact, lead us to draw spurious conclusions by imputing substance which does not really exist in the information. Thus, high powered statistical analyses may be mistakenly relied upon for illegitimate purposes. As Glaser and Strauss note: "They are merely techniques for arriving at a type of fact. It is still up to the analyst to discover and analyze the theoretical relevances of these facts."[16] Quantitative techniques must not be adopted wholesale or in part without a thorough consideration of their implications in any particular investigatory situation. There are, undoubtedly, sets of circumstances within which attempts at quantification will obscure the real meaning of events and frustrate our investigative purposes.[17] It is certainly arguable that in many cases where the approach does render accurate and consistent data, that information may be accurately and consistently invalid.

A final point should be clarified here. It has been noted that the stated workshop goals presented in Study I deal in highly complex and abstract cognitive constructs. The "understanding" goal would appear to be particularly

complex and abstract. Our understanding of "understanding" and our ability to measure its dimensions are rather limited. It is these qualities of abstraction and complexity that make the attainment of such goals so difficult to measure. Had workshop goals been much less complex and abstract, it is unlikely that such a large number of specific objectives would have been required to comprehensively represent them. Furthermore, it is likely that specific objectives could have been more easily defined in operational terms. Simpler goals might, therefore, have meant that the classico-experimental study would have been a more practical undertaking. The drain on resources required to gather information might have been cut substantially while a high degree of validity might still have been maintained. Thus, one may conclude that in instances where valid instructional objectives can be defined in operational terms which do not require large investments of resources for measurement, the classico-experimental paradigm could provide a practical approach to the production of valid and reliable information.

Generalizability

A fourth question which might be asked about Study I and the classico-experimental approach is, "To what extent can valid generalizations be drawn from the information produced?" Since the information derived from the performance test in Study I does not appear to be very reliable, it would probably be foolish to generalize from it at all.

If the same performance test was used again under almost identical circumstances, it seems entirely possible that a different set of information would eventuate. Unreliable data does not lend itself to valid generalization. But what if the information produced did appear to be highly reliable? Could valid and useful generalizations then be made?

In our short discussion of generalizability in Chapter I we reached the conclusion that the validity of generalizations depends upon the extent to which our information represents the universe of information to which we wish to generalize. If our information is a highly representative sample from the universe of information which interests us, generalization from it to that universe may be pre-eminently valid. If our information is not a very representative sample of the universe of information which interests us, generalizations will be highly tentative. Assuming that the information derived from the performance test in Study I did appear to be reliable, its content would still be very limited and, therefore, it would be of limited value for generalization. This is not simply because it indicated that no significant effects had been achieved. Even if we had found significant effects in the intended direction, the data's generalizability value would have been relatively low. The representativeness of information produced by the performance test in Study I is limited by: characteristics of the performance test itself (e.g., dimensions of the case studies, dimensions of the

task set, characteristics of the rating procedure), characteristics of the possible outcomes investigated (only those pre-specified by objectives were included), characteristics of the students (personal, social, job-related, educational), characteristics of the instructors (personal, social, educational, etc.), characteristics of the workshop content, characteristics of the instructional methods employed, charactersitics of the physical surroundings, and so on; the list could continue almost indefinitely. Of course information on the workshop produced by <u>any</u> sort of study will be limited by many of these characteristics. The point is that the demands of the classico-experimental approach itself tend to impose a number of limitations on the information collected and that, because of the relatively rigid nature of the paradigm, those limitations are disturbingly extensive. For instance, if the information available to us is concerned only with the achievement of three pre-specified objectives, generalization to a universe of information concerned with a myriad of other possible outcomes is extremely limited. If we are interested in possible effects which the program might have had upon participants' colleagues or upon instructor-practitioner communications, generalizations would have to be highly tentative. By pre-specifying that only a very limited number of possible outcomes will be investigated, the classico-experimental paradigm limits the extent to which valid and useful generalizations can be made from the information produced.

It has also been noted that the classico-experimental paradigm places little emphasis on descriptive information about the circumstances surrounding the instructional program. Without information about physical circumstances, social norms, personal characteristics, communication patterns, and so on, we are left with very few indications about the validity of those generalizations which we do wish to make. Thus, by limiting the investigation to use of certain kinds of information, the paradigm further limits the extent to which we are justified in generalizing from the information collected.

One can, however, see that certain characteristics of the paradigm provide strength to the limited generalizations validly formed from the information produced. If particularly important variables can be stringently defined and controlled, though the resulting information may be rather limited in scope, the generalizations which can be constructed from it acquire strength through precision. In other words, the limiting properties of the approach should force us to take a great deal of care in generalizing. For instance, by clearly specifying the outcomes to be examined, we clearly specify a range of generalization which is not warranted by the data. Further, those generalizations which do appear to be warranted, though rather limited, are imbued with very precise definitions.

In some ways, the classico-experimental paradigm appears to represent a "safer" approach to the production of validly generalizable information. The likelihood of

inappropriately generalizing from unreliable data should be reduced if the data's reliability can be readily assessed. Furthermore, if important variables can be stringently defined and controlled, generalizations from the data will gain strength through precision. At the same time, the precise specification of objectives, means of measurement, and the conditions under which measurements are to be taken operate against generalization to all but the most limited of universes. Thus, it would appear that a tension exists between the "safety" and precision of quantified methods and the possibility of much broader utility associated with less precise methods of observation. On the one side, Glaser and Strauss observe:

> "Facts" change quickly, and precise quantitative approaches . . . typically yield too few general concepts and relations between concepts, to be of broad practical use in coping with the complex interplay of forces characteristic of a substantive area.18

From the other point of view, Breedlove notes:

> If the results of an evaluative study are to be generalized to another situation, it is necessary to establish convincingly that the new situation, including the new treatment, is comparable to the one in which the research was conducted. The assurance that the new situation replicates the study situation depends upon the extent to which the treatment or intervention variables can be replicated. Whether this can be done depends upon the care with which the variables are identified and measured.19

Before we move on, reiteration of our primary function as evaluators suggests that the foregoing arguments may not be quite as important as they would initially appear to be. It should be remembered that our primary concern is with the <u>evaluation</u> of a particular program of instruction. We

wish to produce high quality information upon which decisions about the instructional program (especially "course improvement" decisions) can be made. Our major concern is not with research on dimensions of instruction or learning as such. This means that, although we will wish to make certain kinds of generalizations--e.g., in order to make decisions about "exporting" programs to other sets of circumstances--most of the information required is of a different order. Those who perform research on "learning" or "instructional methods" usually wish to produce valid generalizations about relatively broad universes of information. Our primary evaluative intention is considerably more restricted than that. If evaluation is seen as applied research, primarily aimed at informing decision-makers who wish to affect program improvement, the focus must be on meaningful and useful feedback at the local level.

Utility

This leads directly to a closer examination of the utility dimension. It was suggested in Chapter I that the utility of information should be assessed against the expense required in obtaining it. No matter how reliable, valid or generalizable the information might be, if obtaining it will require the expenditure of more resources than are available, the information medium is of little practical use to us. Some of our discussion has already indicated that the classico-experimental paradigm may be an impractical approach to meeting many of our most pressing information needs.

The procedures followed in the performance test segment of Study I appear to have been quite expensive, especially in terms of the time and effort involved. In all, fourteen distinct steps were necessary:

(1) the dimensions of the performance test were conceptualized, including the discussion of objectives with tutors;

(2) the case studies, instruction sheets, and explanatory letters were composed, typewritten, and duplicated;

(3) "exercises" were randomly assigned to participants and records begun;

(4) pre-tests were mailed out to subjects along with the letter of explanation;

(5) subjects completed the assigned task and I collected responses upon their arrival at the workshop centre;

(6) post-tests were mailed out;

(7) subjects completed post-tests and mailed responses to me;

(8) thirty-two sets of "notes" were rewritten in standard form;

(9) tutors were convinced to act as judges;

(10) rating tables and instructions were composed and mailed to tutors along with the "notes";

(11) tutors read case studies, assessed responses and posted their ratings to me;

(12) statistical compilations and tests were carried out on the data;

(13) interpretations and conclusions were drawn;

(14) records were maintained throughout.

It would be extremely difficult to specify the exact amount of time and effort that went into this undertaking but it is estimated that, between evaluator, subjects and judges, a minimum of 200 man/hours of work took place--excluding "write-up" time. Financial expenditures, though perhaps not high in "absolute" terms, were somewhat higher than those incurred in Study II; especially at Step 2.

In return for this investment, experience in Study I suggests that the information rendered may not be worth the expense of obtaining it--depending upon our evaluative purposes. On the positive side, the classico-experimental paradigm: gathers very little "unnecessary" information, produces data in a form which promotes ease of handling, and places substantial emphasis on methodological quality; especially reliability. Some of the observed deficiencies in the paradigm, however, suggest that other factors may be of even greater importance in terms of utility. Based on follow-up of twenty American evaluations, Patton found that methodological quality is not a major factor in explaining the utilization or non-utilization of evaluation results.

> Social scientists may lament this situation and may well feel that the methodology of evaluation research ought to be of high quality for value reasons, i.e., because poor quality studies ought not be used. But there is little in our data to suggest that improving methodological quality in and of itself will have much effect on increasing the utilization of evaluation research.[20]

On the other hand, it would appear that relevance-- ". . . the extent to which the information provided in an evaluation

relates to the original purposes of an evaluation"--is of importance to evaluative utility.[21] Patton found that: ". . . the most valuable information with the highest potential for utilization is that information that directly answers the questions of the individual(s) identified as the relevant decision-maker(s)."[22]

If our major evaluative purpose is to produce information which will be valuable to decision-makers who are primarily concerned with program improvement, the classico-experimental paradigm may not prove to be particularly useful. For example, it seldom provides much indication about those elements in an effective instructional program which tend to support its overall effectiveness and those which do not. It seldom suggests how changes in one component of the program will affect other components; or program effectiveness, or efficiency. It usually gives us no information about problems and issues in program implementation; muchless does it suggest solutions or resolutions. In short, no matter how well the study is conducted and how high the methodological quality, the paradigm furnishes us only with information useful to validation. If we have need of more than a validation study, the paradigm's deficiencies may well outweigh advantages like ease of data handling and precision in data gathering.

In summary, one might conclude that the classico-experimental paradigm, when used to evaluate instructional programs involving rather complex learning behaviors in students, is expensive and produces information that falls

short of meeting many formative evaluation needs. If expenses can be reduced through the development and use of better devices to measure complex learning behavior, its utility is likely to increase. However, by its very nature, it will remain useful only in assessing the effectiveness of an instructional program as defined by pre-specified objectives. Although validation studies are an important component of most evaluation, they certainly do not comprise its whole. If program improvement is our major evaluative goal, the paradigm may represent a rather ineffective approach.

Student Ratings

Student rating scales, as used in Study I, do not really fall within the purview of the classico-experimental paradigm. It was, however, anticipated that some form of supplementary information would be desirable and that student ratings might combine with performance test information to produce a more useful evaluation report. It was further hoped that the ratings, in themselves, might afford substantial quantities of useful information. Since rating scales are very often used in the evaluation of instruction, and since Study I offered the opportunity to use them without prejudicing the requirements and outcomes associated with the classico-experimental approach, it was decided that at least a superficial examination of the medium would be included in this investigation. I will end this Chapter therefore, with a few somewhat fortuitous observations and conclusions about the quality of such data. I hasten to

caution the reader that this discussion should in no way be construed as a thorough examination of the subject; simply as a rather casual addition to the main body of the enquiry.

As is the case with information rendered by pre-specified performance tests, the validity of student rating information is limited by the pre-specification of subjects to be investigated. Generally, rating scales must be constructed before the instructional program begins if they are to be applied during its implementation. It may, therefore, be very difficult or even impossible to include items suggested by student-held goals. However, because a pre-test is usually not used, it may more often be possible to postpone the construction of rating instruments until after the instructional sequence has commenced and the inquiry into students' goals has been undertaken. Again, serious impracticalities can arise if extensive validation research on the instrument is to be completed prior to application. In the case of Study I, however, both the "Daily Evaluation" and "Overall Evaluation" instruments were developed before workshop onset, and no research effort was made to validate them. The "Daily Evaluation" forms were to be applied on every workshop day and the first "Overall Evaluation" form was to be applied before participants left the instructional centre. Since only one investigator was available, it was decided that enquiry into student-held goals and construction of a rating form while the workshop was in progress would be an impractical plan of action. In the same way, no serious

attempt at instrument validation was undertaken for lack of resources. Circumstances very often make such preparatory research impractical, but student rating instruments can still be found to have value.

The validity of student rating scales would appear to be enhanced by the fact that quite a large number of diverse items can be included rather easily, thereby encouraging comprehensiveness in assessment. In Study I, for instance, students assessed their small group experiences four times on nine dimensions, as well as assessing the workshop experience as a whole on fifteen dimensions. This provided opportunities for evaluation with regard to quite a variety of workshop elements.

If the objectives of an instructional program involve the encouragement of certain emotive states, impressions, or beliefs in students, student rating instruments would appear to offer potentially high content validity. If, for instance, we wish to assess the extent of a student's belief that he has learned something, presenting him with rating items which concern "amount learned" would appear to lead us toward the acquisition of valid information. His subjective impressions are precisely what we wish to measure. However, if we wish to assess a student's actual learning, the rating item may not provide particularly valid information. For a number of reasons, a student's reported subjective impressions may be an invalid representation of his learning. The subject may report his impressions inaccurately: perhaps to impress others, perhaps to protect

himself from criticism, perhaps simply to justify behavior. Alternatively, he may accurately report erroneous conclusions about the amount he has learned. Obviously one should be rather wary of the assumption that self-reporting of learning achievement represents a valid or reliable gauge of the actual amount learned.

Doyle reports on a number of studies which relate student ratings of their instructors to "objective" measures of learning, and reaches the following conclusion.

> There is no acceptable basis in these studies for accepting the hypothesis of a negative relationship between ratings of over-all instructor ability or effectiveness and learning-oriented criterion measures. On the other hand, neither are the data persuasive that there is a very large positive relationship. There does, however, appear to be a fairly consistent low-to-moderate positive correlation between general ratings and student learning.23

In general, the reliability of ratings decreases as the number of scale points employed decreases. The more scale points, the more distinctions that can be made. However, the inclusion of too many points on a scale can irritate and confuse the rater. For this reason, most of the scales on rating instruments used in Study I contained six or seven points; probably an optimal number.[24] By using a line scale format for most items, the possibility of making finer distinctions was left open to the raters. A number of other precautionary tactics were also employed in an effort to minimize the occurrence of errors. For example, to combat the tendency toward "automatic reponses", the direction of anchors was alternated occasionally. Some raters tend to respond rather "automatically", leaning

toward the marking of each scale in the same general area regardless of the content of the stems. It is argued that this can be prevented by changing the "direction" of the scale occasionally; i.e., by occasionally switching the "favorable" and "unfavorable" ends. Naturally these alterations will require that the rater give greater concentration to the rating task, and they may tend to confuse raters. For this reason, the instructions included an explicit warning that all scales did not run in the same direction and requested that special care be taken. Efforts were also made to isolate "dimensions" from each other so that no interdependence would be assumed. For instance, the five items referring to "facilities" were grouped under one item heading. Although this might tend to promote a certain amount of error within the five item set, it also tends to isolate "facilities" from all other dimensions, thus decreasing the likelihood of contamination in and from other items. Finally, one notes the initial assurance with regard to confidentiality. This was included in the expectation that a relatively firm statement with regard to anonymity would not only encourage individuals to complete the rating form, but to complete it as accurately as possible.

Our discussions in Chapter I warned us of other rating errors which might occur without being at all obvious. "Halo effect"--the rater's tendency to allow overall impressions to influence ratings on specific traits or outcomes--blurs distinctions. The student may, for instance, allow his general respect for an instructor's expertise to influence

his rating of obviously weak teaching behavior. This is not the same as "leniency effect", where the rater tends toward generousity. "Halo effect" can operate in either direction. On the "Daily Evaluation" scales, for example, a rater's response to Item 1c ("relevance to your needs") might be influenced by undifferentiated feelings about the session being rated. If he had found the session to be a particularly instructive experience, he might well have rated it high on the "relevance" scale without really considering specific aspects of relevance. A similar sort of error may occur when the meaning of one item is confused with the meaning of an item or items preceding it ("proximity error"). A rater might, for example, respond to Item 2 ("learning") on the "Overall Evaluation" instrument as a continuation of his response to Item 1 ("enjoyment"). A number of other somewhat similar errors tend to further reduce the reliability of rating data. The subject may rate items similarly out of a feeling that they somehow ought to go together ("logical error"), he may project some of his own perceived deficiencies onto a ratee ("contrast error"), he may tend to rate toward the midpoint of a scale ("central tendency"), and so on. To the extent that such errors are due to individual "chance" variations, they are virtually unavoidable but amenable to statistical correction. To the extent that they are reactions to instrument deficiencies and failures, they may be avoidable.

Although no extensive reliability assessments were undertaken on the rating data produced by Study I, the fact

that the "Overall Evaluation" was administered on three occasions permits a few conclusions. The very high levels of correlation between instances of administration provide strong testimony to the instrument's stability (retest reliability). Applications I and II (one month apart) correlate with each other very significantly, especially on relatively "tangible" sorts of items (e.g., accommodation and meeting space). Correlations between Applications II and III also reach statistically significant levels, indicating that approximately the same phenomena were being assessed in approximately the same ways on all three occasions. The overall decline in ratings (statistically significant between Applications I and III) can, therefore, be looked upon as a relatively reliable representation of a real decline in participants' favorable impressions of the workshop over time. The instrument's apparent reliability also lends strength to a conclusion that the indications of overall favorability of results are reliable.

In Chapter I we made mention of an array of studies which have some bearing on the generalizability of student rating data. Comparisons between student ratings and instructor self-ratings indicate a general tendency for students to rate instructors less favorably than the instructors rate themselves; but students and instructors tend to agree about the pattern of the instructor's teaching strengths and weaknesses.[25] Comparisons of student ratings with alumni, colleague, and administrator ratings of instructors suggest that, while there certainly are similarities,

they are generally too diverse to warrant generalization from one to another.[26] Other studies have looked into the influence of factors such as demographics, ability, and motivation on student ratings. These tend to demonstrate that few such student-linked characteristics influence student ratings significantly.[27] Social factors, such as the instructor's presence and demeanor during the rating period, and the student-perceived social uses to which data will be put, do seem to influence ratings.[28] On the whole it would appear that there are a number of circumstances surrounding the instructional program and the evaluation itself, including the source of data, which do influence ratings differentially. To the extent that such factors suggest limitations on the information's representativeness of universes of interesting information, they suggest limitations on the generalizability of the information.

In Study I there would appear to be little reason to believe that the ratings obtained would not be replicated by almost any sample of professional social workers who received similar instruction under similar circumstances. There was not a great deal of variance in the rating data--indicating that variance in individual characteristics was not a particularly strong source of influence on the results. This bodes well for the generalizability of the information across student groups; generally a most important dimension when considering "course improvement" decisions.

Doyle argues that student ratings should be used in assessing the "instructional component" of personnel

decisions. He suggests that students have greater opportunity than anyone else to view instructional sequences and that student ratings are usually more reliable than--and at least as valid as--any other measures of classroom performance.[29] Thus, when personnel decisions must be made, student ratings should be primary data upon which to base them. He indicates that probably the most productive approach is to ask students straightforward questions like, "How would you rate the general teaching ability of this instructor?" Such an item is valid to the extent that it is reliable (i.e., it means what it says) and infers quite a number of qualities and outcomes commonly considered appropriate in assessments of teaching ability. On other instructional dimensions of personnel decisions, like "choice of objectives" and "extent of relevant knowledge", colleagues' and administrators' ratings might also be used. The opportunity to observe behavior related to these dimensions is relatively available to them as well. Though we might assume colleagues to be in a better position to assess the extent of an instructor's relevant knowledge, however, there is no hard evidence to support the assumption. Finally, one must remember that it is the instructor's methods, techniques, and abilities which are to be evaluated; not the instructor as a person. Further, the instructor is not, and cannot be, responsible for a very great deal of instructional outcome. Since he has so little control over the numerous and diverse forces which may influence outcomes (individual student learning characteristics,

many institutional arrangements and physical circumstances, availability or scarce resources), it is both incorrect and unfair to hold him totally responsible for outcomes.

Investigations have demonstrated that, under the right conditions, student rating feedback can and will produce changes in instructor behavior.[30] Since the research literature provides little basis for assessing the appropriateness of such changes, however, we must try to ensure that the wording of rating items reflects sound judgements about what constitutes good teaching in a particular set of circumstances.[31] For, the validity of student ratings is largely determined by the appropriateness of the words used to construct them. Unfortunately, the breadth of rating items often operates against acquisition of precisely the information we need. That is to say, because of their breadth, rating items may not give us the specific information we need to decide upon appropriate alterations to an instructional program. For example, "Overall Evaluation" Item 3 has validity to the extent that its words say what we intend them to say: "How relevant was the workshop content to your on-the-job needs?" It would appear to pose a relatively straightforward and unambiguous question. The item's very breadth, however, precludes collection of information which would tell us about the specific areas of content that were perceived to be especially relevant and those which were seen to be most irrelevant. Responses to items like this, then, may provide indications with regard to broad areas of strength and weakness in the instructional

system, but they do not provide a great deal of assistance in discovering specific guidelines for remedial action.

It might also be noted that the rating instruments in Study I provided a means of gathering some information on the applicability or utility of workshop content to on-the-job needs. By use of two follow-up applications of the "Overall Evaluation" forms, subjects were given a medium through which they could feed back some evaluative information after attempting to apply a unitary approach on the job for a number of months. Further, since the instrument was already constructed, the cost of obtaining such information remained relatively low. Although it does suffer from some of the limitations already described, follow-up information of this sort may be very valuable--particularly in an instance like this where instrument stability appears to be quite high. For it is in the more long-term effects that objective achievement is most desirable. Indications of applicability to on-the-job conditions--or the lack of it--provide information of real value to both course improvement decision-making and the accountability function.

In Study I the existence of student rating data afforded a check on the reliability of performance test information. In combination with performance test data, they also offered an opportunity to use quantitative data in answering some possibly useful questions (e.g., Did those who apparently developed greater "understanding" perceive certain workshop characteristics to be especially

valuable?). In themselves, ratings gave us a set of relatively valid and probably reliable information upon which to base evaluative decision-making. Although the data's utility is bound by certain limitations, it seems a justifiable conclusion that, especially in combination with other sorts of data, the considered use of student rating scales significantly contributes to the accumulation of information useful to evaluation.

We would be remiss, however, in failing to note that obtaining this information may necessitate relatively high expenses. The construction and application of student rating scales can require heavy expenditures of time, effort and money--especially if extensive validation research is carried out prior to their use. In Study I the financial expenditure was reduced by application of the first "Overall Evaluation" along with the performance post-test. However, the other two applications and the voluminous printing task required substantial expenditures, construction of instruments required more effort than was originally anticipated, the "Daily Evaluation" applications were somewhat disruptive of workshop routine, and data compilation and analysis was a tedious and time-consuming task. Perhaps the relatively high expenses and somewhat limited nature of the data obtained, suggest that student rating instruments are <u>best</u> used in combination with other sorts of information-producing media; particularly when program improvement decisions are our primary intent.

CHAPTER VII

"STUDY II AND THE SOCIO-ANTHROPOLOGICAL PARADIGM"

Since the socio-anthropological paradigm advocates utilizing a range of investigatory media to examine a variety of practices, issues, questions and outcomes associated with an instructional program, its dictates are not so intransigent in application as those of the classico-experimental approach. The scope and flexibility afforded to the investigation by the paradigm render it somewhat more practical and easy to apply. We noted that a major limitation in using the classico-experimental paradigm resulted from its inflexibility in data gathering. It does not appear to be well suited to investigating, or adapting investigations to, significant changes in the instructional program or its environment while the study is in progress. For various reasons instructional systems will naturally change in a variety of ways, and an overall conceptual framework for evaluation must provide a capability for coping with the associated intermittent decisions and alterations in direction. Through the use of a variety of data gathering media and data sources, as well as the principle of progressive focusing, the socio-anthropological approach appears to be much more flexible in coping with ongoing program and contextual change. Information sources or media used can be tailored to the specific circumstances and requirements for evidence that arise. New data collection techniques, foci of attention, and objectives can be

appropriately introduced at any time; controlled only by characteristics of the emergent phenomena, available resources, and the purposes of the study. In Study II, for example, attention was drawn to the use of participant-generated case material as it became obvious that students in the small groups were rejecting the prescribed case studies and turning to consideration of more "realistic" material generated from their own working experiences. It was because investigation began as a generalized inquiry into the processes and structure of the workshop, and progressively focused on significant events and relationships, that this unanticipated change in the material used was noted and followed-up.

This flexibility, however does not mean that the approach is without demands and prescripts. One such ordinance arises from the emphasis which the paradigm places on the evaluator's initial entry into the field of study with few preconceptions about what he will find there. It is intended that a minimum of preconceptualization will permit the investigator to less prejudicially examine the myriad dimensions of the instructional system as they exist, rather than as he might otherwise expect to see them. For precisely that reason, Study II was actually undertaken before Study I. Even though very different evaluation paradigms were used, it was anticipated that involvement in Study I might evoke preconceptions about the Unitary Approach Workshop which could colour my interpretation of events in Study II. The socio-anthropological

approach relies heavily upon the observations and subjective interpretations of the evaluator. If I had attempted to apply that paradigm after observing and examining an extremely similar instructional sequence from the classico-experimental viewpoint, "contamination" in my mind between the two programs was likely to have occurred. Since the classico-experimental approach stresses utilization of "objective" measurements and de-emphasizes subjectivity on the part of the evaluator, the "purity" of my own impressions was not so important in its application. Thus, in applying the socio-anthropological paradigm first, an initial effort was made to satisfy the prescripts of that approach to evaluation. I did not, however, begin Study II entirely dispassionately. I undoubtedly possessed a number of preconceptions about the program; preconceptions which derived from a variety of sources: the discussion with workshop tutors about their aims and the proposed program in general; study of the reading kit posted to prospective participants prior to attendance; some reading of the unitary approach literature; basic information about characteristics of tutors and attendees; and, last but certainly just as important, memories from a number of rather similar instructional programs which I had either attended or instructed (I had, in fact, previously attended one week-long seminar with very similar content to the workshop). So, though I attempted to retain as much objectivity as possible, I did suppose quite a great deal about the workshop before it even began. We have already noted that proponents of

the socio-anthropological approach to evaluation do not suggest that the evaluator attempt to enter his field of study with a completely "blank mind"; simply that he should begin with ". . . only general research questions, in chapter-heading terms, and nothing like the orderly research design of his counterpart in the other paradigm."[1] The general research questions with which I entered Study II revolved around three seemingly important issues: the utility of the small group-plenary session format, the teacher-practitioner configuration of the participant group, and the theory-application orientation of the tutors. However, other than the "Participant Information Sheet" which was common to both Studies, I began the investigation with no preconstructed observational instruments; only pens, paper and a tape recorder. It was my intention that the unfolding course of the workshop should determine the course of my Study.

In order to further examine the extent to which Study II conforms to the principles and ordinances of the second paradigm, I have extracted five apparently important prescriptive ingredients underlined by the proponents of the socio-anthropological approach: (1) focus on "learning milieu", (2) production of descriptive data, (3) "progressive focusing", (4) variety of data gathering methods and "triangulation", and (5) interpretation and explanation. These will be used as bases for ordering this first part of our examination.

Focus on Learning Milieu

Even a cursory reading of Study II will disclose its preoccupation with examining the context within which the workshop actually operated and the instructional processes which occurred to create that learning milieu. At the same time, the Study almost entirely disregards the "unitary approach" content of the workshop and certainly does not focus attention on the "measurement of outcomes". The lack of attention to substantive content did not directly result from any injunction of the socio-anthropological approach; the paradigm does not provide any such counsel. As was the case in Study I, I found that restraints imposed on the resources available forced me to circumscribe the scope of the Study. With only my own perceptive apparatus, skills, financial resources, time, and efforts to draw upon, I found it impossible to adequately attend to the context and processes of interaction occurring while simultaneously maintaining a record of the content which emerged. Since the paradigm appears to place greater stress on the former, that is where I concentrated my limited resources. A more thorough illuminative study, however, might have comprehensively related workshop content with "learning milieu". This is, therefore, one area in which Study II falls short of being an ideal application of the paradigm. Be that as it may, the focus on learning environment which was taken creditably conforms to the injunctions of the paradigm. Through direct observation, interview, discussion with participants, and application of a questionnaire, an attempt

was made to create a kinetic portrait which reflected the most notable elements of the workshop as a learning milieu.

Descriptive Data

Portraiture denotes description and a large proportion of the information gathered in Study II comprises a highly descriptive account of the workshop: its participants, interactions between participants, reported intrapsychic states, physical circumstances, interaction between participants and material, and so on. The central portion of the report is entirely narrative. Even the interpretative and conclusional comments attempt to convey "impressions" radiating from the various points of view discovered. If nothing else, the Study was descriptive.

It might be noted here that the effort to generate so much descriptive information was substantially, if inadvertently, aided by my working arrangement with the workshop tutors. They had requested that I refrain from actively participating in workshop sessions, or from consciously providing evaluative feedback to them while the workshop was in progress. They wished to minimize the amount of influence which my presence would have on the workings of the instructional system. Since my initial intention had been toward a more "formative" sort of evaluation, I had mixed feelings about the stipulation. It did, however, allow me to more completely focus my attention on observing and recording details of workshop events as they occurred.

Progressive Focusing

My initial aim, upon entry, was to familiarize myself as thoroughly as possible with all apparently significant aspects of the workshop milieu. As time went on, particular problems, issues and practices were certain to emerge frequently enough to warrant closer attention. In the event, this is precisely what did happen. For instance, the concern with a need for more didactic input arose repeatedly as it became increasingly obvious to participants that little would be forthcoming. By the time the senior tutor's lecture on elements of unitary approaches was complete, it was clear that some participants were confused and wanting more directly-presented information. The "analysis of a system" exercise did not instil sufficient understanding for their liking and, by the time that session was completed, numerous verbal expressions of the unmet need were heard. In similar manner, the utility of participant-generated case material became increasingly obvious as the workshop progressed. Individuals who reported with the greatest enthusiasm about learning from their small group experiences often mentioned this factor as a positive influence. Similarly, the controversy over the decision by Group B members to remain together rather than reconstitute highlighted a felt need on the part of many workshop members for closer, more open communication with tutors about the ways in which the workshop was proceeding. None of these issues were preconceived by me as such; they arose from immersion in the ongoing life of the instructional

sequence. It is possible, however, that I began to see them developing as common issues before many participants did because I was in a better position to notice developments in the workshop as a whole. When an issue surfaced over and over again, repeatedly referred to by different individuals from different working groups, it rather quickly took on an aura of importance.

Other areas of concern--such as the benefits of group composition, problems in the application of relatively abstract concepts (especially "systems" concepts), tensions between social work teachers and practitioners, and time problems in task completion--were anticipated very early in the workshop. Some were, in fact, suggested by my "pre-knowledge" of the workshop design. On the other hand, I also anticipated that there might be serious problems with termination, scheduling, and meeting space which, in the event, did not turn out to be particularly consuming. By remaining deeply involved in the instructional sequence, spending all of my time observing instructional processes and talking with participants, my attention was quite naturally directed toward those problems, concerns, issues and notable practices which occurred with the greatest frequency and force. I could see, for example, that the concern with small group composition was arising in the group sessions, in leisure-time chat, and in the tutors' meetings. It therefore naturally became an area which required some further investigation. The paradigmic prescript toward progressive focusing, then, was seen to

be consonant with my natural responses to intensive investigatory involvement in the learning milieu.

Data Gathering and "Triangulation"

Quite a variety of data gathering media and data sources were used in Study II. The written program, documents provided by the tutors, and responses to the NISW's short evaluation form were examined. The "Participant Information Sheet" collected background data on personal, social, professional and educational characteristics of workshop members, and also provided for an initial assessment of their expectations and motivations prior to workshop attendance. I spent a substantial amount of time and effort in first-hand observation of workshop sessions, meetings, and informal gatherings; as well as in informal conversation with individual participants and tutors. Relatively extensive interviewing was conducted with each of seven participants, and an open-ended questionnaire was used to gather "follow-up" information from most participants approximately five months after the workshop. The "principle of triangulation" was consciously applied so that each issue, concern, or problem would be viewed from more than one perspective. For instance, difficulties encountered in attempting to apply "unitary approach theory" were first noted in observation of the small group sessions, they were inquired after in the interviews, they arose quite naturally in informal conversations, and they were further examined via the follow-up questionnaire. The question of participants' motivations to attend was focused on in the

participant information sheet, investigated through observation of the first small group session, and inquired after in interviews as well as informal conversations. In fact, every issue or concern which appeared to reflect possibly significant elements in the learning milieu was investigated from the perspectives provided by more than one medium.

Interpretation and Explanation

Perhaps it is in the area of "interpretation and explanation" that Study II is most subject to criticism. The report included generalizations, inferences, and interpretations with regard to a number of issues and concerns. It attempted to support them with information gathered through all media and from various sources. To the extent that relatively distinct patterns could be discerned in the mass of information collected, they were presented in the report. However, some pieces of information tended to contradict others. It was found that while the media tended to support each other, sometimes the sources did not. Only in relatively few areas (e.g., benefit of interaction with peers, need for more didactic input, need for greater heterogeneity in membership) did respondents tend to agree almost unanimously. In most areas (e.g., program scheduling, small group reconstitution, "outstanding" sessions), even if a relatively common viewpoint was expressed, a variety of dissenting voices was heard. Consequently, in order to report on dissenting opinions and observations, interpretations and explanations had to be radically tempered. Furthermore, much generalization beyond the context of

this particular type of instructional program did not appear to be warranted; it was felt that the data would not support generalization into somewhat dissimilar realms of instructional endeavour. This limitation will receive closer attention later in the Chapter. For now it is simply noted that the interpretations and explanations put forward in Study II lack a degree of force, primarily because contraindications were almost always present in the data.

Our short examination of these five paradigmic prescripts, as applied in Study II, suggests that the Study does represent a serious attempt to apply the socio-anthropological approach in evaluating a short course for social workers. It concentrated investigations on the context and processes which went to make up the "learning milieu". It provided much descriptive data. It quite "naturally" allowed emerging trends, issues and concerns to guide "progressive focusing". It used a variety of data gathering media and sources to "triangulate" on emergent issues. It attempted to discern patterns in the information which would suggest interpretations and explanations. The socio-anthropological paradigm appears to be conceived broadly enough that almost any data-producing medium can be appropriately utilized in relation to the problems, concerns, issues and potencies which emerge from the learning milieu. Since no rigidly prescribed set of discrete "steps" is advocated, research methods can be adapted to the practical limitations imposed by the instructional program and adverse

circumstances surrounding the investigation. Furthermore, the use of almost any investigative technique can be justified as consonant with the demands of the paradigm, so long as none are used in isolation. Thus, it is not difficult to demonstrate that the methods of inquiry used in Study II combine to represnt a valid reflection of paradigmic ordinances.

We will now turn to examination of the quality of information produced by the data-gathering media used in Study II. Information resulting from direct observation, discussion, interview, examination of documents, and administration of questionnaires will be assessed in terms of the four "quality" dimensions outlined in Chapter I. As in the preceding Chapter an effort will also be made to extrapolate from our examination of Study II in order to shed more light on some of the advantages and limitations associated with the socio-anthropological paradigm itself.

Validity

Both validation and evaluation require fact and judgement. Decision-makers must not only consider information about elements in an instructional sequence and the program's notable consequences, but must also consider them in the light of associated value orientations. For example, an instructional program which evidences no significant educational outcomes may, nevertheless, be considered desirable. A program which is demonstrated to be highly effective may, nonetheless, be undesirable. Value dimensions are relevant to decision-making. The qualitatively-

based socio-anthropological approach to investigation can quite explicitly include considerations of value dimensions; particularly of the values held by program participants. The decision-maker's understanding, therefore, need not be restricted to that which he can garner from "hard" evidence which is not explicitly value-related. Questions of human motivation, intention, dislike and priorization will be integral to a socio-anthropological inquiry because such an approach conceives of social and psychological phenomena as essentially human products. As Leonard says, such an approach ". . . has the important effect of restoring attention to the subjective world of the individual and thus helps us to understand more fully his interaction with the social processes around him."[2] Thus, where the classico-experimental paradigm was apparently not well suited to evaluating with regard to student-held objectives, the socio-anthropological approach appears to place a good deal of emphasis on participant values--all participants' values. It aims at assisting us to understand relationships between human values, activity, and consequences; and it permits us to inquire more thoroughly into students' intentions.

If one places a great deal of stress on empirical validation, then the validity of information produced by the socio-anthropological approach is most certainly suspect. If, on the other hand, subjective human judgement is endowed with trust, the information rendered by use of the paradigm might be judged to have very high validity. This is a

decision about axiomatic alternatives and requires a value judgement. From the second point of view, the subjective judgement of the investigator himself assumes a great deal of importance. The socio-anthropological approach obviously encourages use of "interpretative human insight and skill".[3] It may be argued that the skilled human mind is by far the best "instrument" available for the collection of valid information about individual human behavior and interpersonal interaction. The mind is able to sift and sort extremely complex arrays of indicators, is attune to nuances and ranges of meaning that no other instrument could register, and is able to make the closest of distinctions when weighing up evidence. If such a position is taken, as it is in the socio-anthropological paradigm, the relevant competence of the investigator becomes a very important consideration. His educational and experiential background, his lack of bias or prejudice, as well as the circumstances surrounding his investigation, must be viewed as significant factors in assessing the validity of the information he produces. One must ask questions like: "Does this investigator have sufficient and appropriate background knowledge in the required subject area(s) to permit valid judgements?", "Did he have sufficient opportunity to view the instructional program from a variety of perspectives?", "Are his judgements supported by tenable explanations of how they were reached?". If affirmative answers can be given to such questions, the validity of information rests heavily on the reliability of the investigator's observations and the validity of other data sources.

The validity of information in Study II was undoubtedly enhanced by the latitude in media and source utilization permitted by the socio-anthropological paradigm. Using a range of information-gathering media permitted me to base interpretations upon data gathered from a variety of perspectives and sources. I was not limited to concentrating attention on measuring the attainment of a few pre-specified objectives. I was free to emphasize the use of data sources or media which appeared to lend the most valid information under the particular circumstances of that instructional sequence. The opportunity to progressively focus on concerns and issues which took on new measures of importance as the workshop progressed, made it possible to collect information which was most likely to be required in tutors' evaluative decision-making. At the same time, the flexibility of the approach permitted a degree of comprehensiveness in data gathering that would not have been obtained with a more rigidly specified set of investigatory procedures. It would appear, then, that this built-in tolerance for variation in method encouraged the generation of a more valid set of information.

It is rather difficult for me to objectively assess my own fitness as an investigator of the sort required both by the paradigm and the particular system under study. My educational background in sociology and social work, as well as my experience in social work practice and instruction, would appear to provide at least a modicum of skill and knowledge appropriate to the tasks at hand. I was

accustomed to systematically observing and analyzing social phenomena. The study of interpersonal interactions occurring within small groups had always been of particular interest to me and, during my preparation for professional practice, I had concentrated most heavily on social work with small groups. Latterly, much of my practice as a social worker had been conducted in the small group mode and the majority of my forays into social work instruction had utilized small group seminar and workshop formats. The structure of the Unitary Approach Workshop was, therefore, not at all foreign to me. On another front, my training, reading and immersion in the culture of professional social work had rendered me relatively familiar with general issues in social work practice as well as some of the literature on "unitary approaches". As mentioned earlier, I had previously attended a similar week-long instructional program based on the Pincus and Minahan conceptualization of a unitary approach to social work practice. Thus, neither social work practice, social work instruction, small group "processes", nor "unitary approaches" were particularly alien to me. Though I could not be considered an expert in such matters, I believe that my relative facility with many of the practices and constructs utilized in the workshop provided me with invaluable preparation for understanding the learning milieu which developed there. On the other hand, I entered into the Study with little experience in formally evaluating instructional systems and no experience in using the socio-anthropological paradigm

for such an endeavour. Furthermore, my knowledge of educational research methods and the results of research on instruction--as well as my comprehension of "learning theories"--were far from impressive. In terms of educational research, I was a novice.

As intimated earlier, one could convincingly argue that my perceptions and interpretations of workshop events were seriously affected by my own preconceptions, attitudes, and knowledge. This, it would appear, is unavoidable if such an approach to evaluation is used. In fact, one might argue that it is unavoidable in using any approach to evaluation. Certainly the use of direct observation, interview and informal conversation to gather information encourages a degree of personal interpretation in the data. It is for this reason that the "orientation" of the evaluator if of special interest when assessing the validity of information obtained through such media. However, the investigator's "orientation" is also important to the extent that it influences the perceptions and behaviors of participants in the instructional process. In the conduct of Study II I was, therefore, concerned with minimizing the influence of my own attitudes and ministrations on participant attitudes and behavior. I wanted to minimize my influence on the data which originated from participants. I was not intended to be part of the instructional "treatment". For this reason, the tutors had requested that I refrain from participating in the formal workshop sessions beyond the point of systematic observation. For the same reason, I

offered very few of my own interpretations and observations when interviewing or simply conversing with participants. This self-imposed control on expression became personally frustrating at times but was unreservedly accepted by participants once they understood the reasons for it. They rapidly came to realize that it was their own observations, impressions and interpretations which were desired, and they respected my attempts to minimize my own influence upon them. On the occasions when I did consciously offer my own interpretations or observations, they were usually as exploratory or tentative statements designed to provoke thoughtful and more specific responses from participants. In a similar manner, the four items included in the follow-up questionnaire were worded in rather general terms and space was left on the form for individuals to respond in their own idiom. It is not argued that these precautions on my part necessarily produced valid data; it may still be cogently argued that my very presence exerted a substantial influence on both the form and content of information tendered. It is, however, suggested that my non-authoritative and non-participative stance did substantially reduce the extent to which my own viewpoints dominated the dimensions of the data acquired.

Obviously, the primary source of information which was external to my own observation resided with workshop participants and tutors. One must, therefore, make some inquiry with regard to the apparent validity of their observations and interpretations. Were these persons

veritable sources of information? Since they possessed ample opportunity to view the workings of the instructional system at first-hand, their observations would appear to be entirely relevant to the investigation. One might also argue that their aims, intentions and desires were of particular importance as criteria against which to judge the effectiveness of the program. As to their competence in observing and interpreting such a program, they were, with very few exceptions, professionally trained and highly experienced social workers, accustomed to observing and analyzing interpersonal events and intra-psychic states. Furthermore, it may be assumed that they were in a better position than anyone else to assess their own intra-psychic states. It would appear, then, that each had acquired a relatively high degree of familiarity with, and understanding of, social work practices and concepts as well as their own idiosyncratic natures. In addition, well over half of the participants and all of the tutors were experienced social work instructors, sensitive to many of the problems and issues which commonly arise in social work instruction. Collectively, they therefore represented an assemblage of relevant experience and expertise of very high order. It should, however, be noted that reliance upon data from a variety if individuals such as these invites a degree of variability in meanings attributed to the words used. Although words do have common meanings, especially amongst a group of individuals with common training and experiences, it is necessary for the investigator to interpret the

meanings implied in each individual's statements and, thus, convert them into a somewhat more regularized form. This may render the information more apparently understandable, but it may inadvertently detract from its validity.

Reliability

Interviews with participants in an instructional system would appear to offer the evaluator opportunities for the collection of potentially reliable, as well as valid, information. Although the information garnered is primarily second-hand, and is therefore coloured by the perceptive set and reporting skills of the respondent as well as by the perceiving and recording apparatus of the investigator, the reliability of such data can be significantly enhanced by the use of an interview schedule which outlines the phenomena upon which most attention is to be concentrated. In Study II, for example, an interview schedule composed of 13 items was constructed with regard to the major issues, concerns, and problems which had surfaced in the course of the workshop. It was loosely followed in each interview. If responses pertaining to these major areas of interest were not "naturally" elicited, an appropriate question was asked directly. On the whole, however, it was found that as interviewees talked about their workshop experience, they tended to cover almost all of the scheduled items with little prompting. This might, in fact, be seen as an incidental indication of the validity of the information required by the schedule; since almost all of respondents' unprompted

expressions were covered by one or more schedule items and since interviewees tended to cover the content of the schedule without prompting, the schedule would appear to have comprehensively represented the issues of greatest common concern. The element of standardization inherent in using an interview schedule of this sort tends to increase the consistency of responses to certain lines of inquiry. All respondents, for example, expressed their views on the group reconstitution and communication-with-tutors issue, on the adequacy of didactic input, and on the utility of the facility in accommodating the workshop. This is not to say that each individual's responses were identical with those of the others, merely that very similar areas of inquiry were covered in each interview. However, though the contents of responses were not identical, there certainly was an air of similarity about them. Many of the same evaluative points were made repeatedly, with no promoting on my part. This similarity of reported observations and interpretations lends a rather strong indication of the information's reliability. Generally, all interview-generated information displayed a very high level of internal consistency. Further, as judged against my own direct observations, it appeared to be very accurate. Since all workshop participants had undergone similar previous educational and working experiences, it is perhaps not surprising to find that they perceived and interpreted phenomena in a relatively common way. Their

mutual heritage of social work training and experience probably contributed to their consistency with each other. Hopefully it enhanced their accuracy as well.

The consistency and accuracy of direct observation data depends upon the perceptual and recording abilities of the observer. Again, reliability may be increased through the use of an observation schedule. It is in the sphere of direct observation that the investigator's knowledge, perceptual sets, biases, and personal motivations probably require the closest scrutiny; the observer himself represents the source and medium of information. In the report on Study II, no effort was made to conceal the inherently subjective nature of the observations and interpretations reported. Although I did endeavor to view events and individuals as objectively as possible, I was ever mindful of the idiosyncratic nature of my observations and attempted to acknowledge them. To the extent that my training, reading and experience has conditioned me to perceiving and thinking from the confines of a limited "set", the information may be biased by that particular orientation. Furthermore, to the extent that my own feelings of liking and disliking coloured my perception and reporting of events, biases are necessarily inherent in the data. One can but report what one perceived, inferred and felt with as much accuracy as possible by trying to ensure that misperceptions, incorrect inferences, and personal projections are kept to a minimum. The information accruing from my direct observations appears to be quite consistent with information gained through the other data-gathering media;

thus, it appears to be relatively reliable. Here, however, is a major issue in using the socio-anthropological paradigm; one must accord with a degree of trust the investigator's integrity, skill and self-awareness. This makes it doubly important that the evaluator consciously acknowledge the subjectivity of his information and attempt to present a picture of the conditions under which it was obtained. He must make a conscious effort to assist the reader of his report in assessing the extent to which information is valid and reliable.

Information gathered from informal conversations is especially open to bias, both from the respondents and the recorder-reporter. In more formally designated evaluation sessions, during interviews, and on questionnaires, respondents tend to keep their personal biases in somewhat greater check; there is a tendency for them to measure their words more carefully before speaking. In the circumstance of an informal conversation, discussion tends to flow more freely; positions are often not so rigorously thought out. Obvious inconsistencies and inaccuracies become more frequent. In Study II, informal conversations were held with as many participants and tutors as possible, and as often as possible. At first, as one might expect, my attentions were received by some participants with a degree of suspicion. They were not entirely sure of my motivations, or of my trustworthiness. Tutors, having established a relationship with me previously and displaying an apparent willingness to help, were initially more forthright in

expressing their opinions and reporting on their observations. Some participants were much more reticent than others. However, as the workshop progressed, participants appeared to become generally more trusting of me and, therefore, more frank. It seemed extremely important that I remain unidentified with any factions or subgroups. In return for appearing to be neither tutor nor participant, neither instructor nor practitioner, neither "pro" nor "con" on any issue, I was privy to relatively confidential and ingenuous intimations. As participants noticed that I did not gossip about other members' reactions and observations, they appeared to become more free in discussing their own.

Although personal biases are more obvious in informal conversations, my experience on Study II suggests that this medium can provide a great deal of very accurate information. Late night discussions in the lounge, short conversations over coffee, and chance meetings in hallways tended to confirm impressions gained through direct observation and interview. Furthermore, such information often sensitized me to the existence of phenomena which I had not previously noticed but which proved to be worthy of later examination. Thus, as long as precautions are taken to guard against outright acceptance of the personal biases which creep in, information received during informal conversations may be invaluable. The most important condition to be observed is that the investigator must not allow himself to treat this information as he would first-hand

material; its veracity must be checked against information collected by other means.

Information gathered from the examination of documents is also primarily second-hand and is, therefore, associated with two sources of unreliability: the provider and the receiver. Although certain documents may serve to acquaint the evaluator with the intentions of course organizers and instructors, they can prove to be rather unreliable sources of information about the program as it actually occurs. This point was noted earlier. In addition, it is worth noting that documents which are produced for reasons other than those reflected by evaluation needs and intentions should be regarded cautiously. Information contained in funding and curriculum proposals may be predisposed to supporting the position of the proposing agency. "Quicky" evaluation forms are often poorly thought out or so "standardized" as to be rendered inapplicable to the particular program under investigation. The documents examined in connection with Study II were not, however, curriculum or funding proposals: rather the written content "summaries" provided as background information for participants prior to workshop attendance, and the NISW's short evaluation form. The evaluation form, though not apparently designed with this instructional program in mind, seemed to be rather well composed. The "background material" appeared to be useful in preparing participants for workshop attendance and did serve to reacquaint me with "unitary approach" concepts. Responses to the NISW evaluation forms, insofar as they were

appropriately related, were used as an additional check on the information obtained by the other media and tended to confirm its reliability. However, neither set of documents was used extensively in the evaluation report. Had the investigation focused more heavily on the content of workshop sessions, the "summaries" and other documents--such as the small groups' "wall-hangings" and reporting notes--might have been used to advantage. As a matter of interest, observation of the groups in process indicated time and again that the visual material displayed in the plenary sessions was not entirely reflective of the discussion content which had occupied members. "Wall-hangings" were frequently concocted in the last few minutes of a group session, apparently more to accommodate the demands of the plenary group situation than to represent the work which had been undertaken. Oral reports in the plenary sessions much more adequately reflected the deliberations and examinations which had occurred.

Questionnaires can be extremely reliable or extremely unreliable information-gathering media; depending, in large part, upon their construction. If sophisticated research is undertaken to assess and improve their reliability and validity, they may yield very high quality data. Since neither my financial resources, time, nor methodological expertise were adequate to preparatory research of that sort, the follow-up questionaire in Study II was kept very simple. It asked four rather generally worded questions with regard to four broad areas of concern and requested that

participants respond in whatever manner they believed to be appropriate. Thus, very few expectations or "cues" were presented and the information contained in each completed questionnaire was largely ordered by the respondent. Although it is entirely possible that the time and space constraints imposed on responses by the questionnaire format--and, perhaps, also respondents' perceived lack of influence on events--might have operated to render questionnaire responses more simple than those gathered in discussions and interviews, analysis suggested a number of patterns or commonly reported opinions and observations which tended to conform with already established conceptual divisions in the data. The information was quite consistent with all other information collected. There was, however, some indication that the five-to-six-month elapsed time period had clouded the participants' memories of various workshop events. Though one might expect workshop phenomena to be viewed in a somewhat different light after five or six months, this did not appear to radically affect participants' responses. It seemed more that some respondents simply did not clearly recall the details which had so preoccupied them a few months previously. They were frequently left with only vague and generalized feelings of enjoyment, challenge, depression, vitality, what have you.

The reliability of the information produced in Study II would appear to be enhanced by the use of a variety of data gathering media and data sources to triangulate on emergent issues. The reliability of information drawn from

one source or with one medium can be checked against information gathered from other practical lines of investigation. The data produced is precisely the sort of evidence upon which people have always made legal, social, and political decisions. In a vast number of instances it is the only form of evidence practically available and, though never unequivocal, the conclusions drawn are often highly acceptable.

Generalizability

We concluded in Chapter I that any piece of information is of limited generalization value; depending upon the extent to which it reflects a representative sample from the "universe of information" to which we wish to generalize. The generalizability of findings from Study II is, of course, limited by many of the same factors which limited the generalizability of information gathered in Study I (e.g., characteristics of the participants, characteristics of workshop content, characteristics of instructional methods used). Certain characteristics of the methods employed to obtain information and the sources from which data were drawn are also limiting factors however, and these are somewhat different in Study II from those in Study I. The characteristics of workshop participants and the evaluator himself would appear to take on special importance as factors limiting the generalizability of Study II information. With the exception of that which was derived from direct observation, the data collected in Study II was all

second-hand information. Someone, usually the workshop members, first made the observations and developed the impressions before they were harvested, compiled and analyzed by the evaluator. The observation and reporting characteristics of participants, therefore, assumed an influential role in the production of the data. This must impose limitations upon the feasibility of valid generalization. The reliability of participant-generated information is not the only component of these limiting effects. The extent to which participant observations reflect a validly representative sample of the information "universe" to which we wish to generalize must also be considered. If we wish to generalize to a universe represented as "continuing education in social work", the observations and interpretations of approximately twenty-eight experienced professional social workers might be considered a validly representative sample of useful information. If we should wish to generalize to the universe represented as "professional continuing education" it is doubtful that the characteristics of our sample would permit more than the most tentative generalizations.

When data is broken down by medium of generation, the generalizability problem becomes even more acute. For example, only seven participants were interviewed in Study II. To what extent can we consider their views to be representative of the views of workshop participants in general, much less those of all social workers or all professionals? Similarly, the information garnered in

conversation with one or two participants over dinner is probably even less validly representative of any hypothetical universe of information to which we might wish to generalize. Thus, it would appear that the small proportion of individuals who can practicably be sampled, combined with the subjective nature of the information, imposes rather severe limitations on the generalizability of the information collected by such methods. Greater generalizability requires a greater infusion of resources and/or the combination of methods.

There is notable variation in the basic orientations to "sampling" which were implied in the two Studies. As we noted earlier, the adequacy of samples used in classico-experimental studies is judged against their degree of representativeness. The adequacy of samples in the socio-anthropological paradigm is judged against how widely and diversely the investigator chose his samples to saturate the categories of educationally significant information desired. Thus, the adequate statistical sample will representatively reflect the characteristics of a population which are considered to be important, while the adequate "focused" sample facilitates comprehensive examinations of instructional issues and aspects judged to be of greatest value or interest. In Study II, the seven participants interviewed were not selected to reflect a statistically representative sample of the workshop membership, but because each had expressed viewpoints and observations which were considered to be instructionally significant. For example,

the workshop's three most vociferous critics were chosen so that their critical observations could be explored more completely. Two of the workshop members who had expressed extreme, not representative, emotional difficulties in coping with small group interaction were interviewed so that the dimensions and possible causes of those difficulties could be examined. Thus, the sample of participants to be interviewed was drawn, not to provide a statistically representative sample of the population, but to enable more comprehensive investigation of those issues and program aspects which appeared to be of particular instructional significance.[4]

Clearly, the adequate statistical sample is useful for generating information from which well-supported statistical inferences can be drawn, while the second form of sample is useful for generating detailed information about chosen topics. Both have value, according to the sorts of information required. Though a carefully drawn statistical sample may not provide us with the desired information in some instances, in other instances assessments of statistical significance will bolster the assurance with which certain conclusions can be drawn. It will sometimes be the case, then, that a study which cannot provide adequate statistical support to its conclusions may be inadequate to decision-making needs.

A similar sort of limitation is evident in information derived from direct observation. To the extent that such information is a product of the subjective interpretation of the observer, its valid generalizability is

limited by characteristics of the observer. We have noted that certain of my own "characteristics" would appear to qualify me as a relatively competent producer of valid information about the Unitary Approach Workshop (i.e., my sociological training, my social work training and experience, etcetera). These same characteristics might not, however, qualify me to validly observe and interpret elements in a rather dissimilar instructional system. For example, my knowledge of computer-based instruction and Saul Alinsky's approach to community organization might well be too limited to qualify me as a competent observer of a program of that sort. Thus, my lack of competence to validly assess this second program imposes another limiting factor on the generalizability of the information produced. To the extent that mine is not considered to be valid judgement on elements comprising a universe of information in which we are interested, generalization must be severely tempered or completely forgone. As we noted in the last Chapter, however, evaluation to effect improvement must, in the first instance, focus on reapplications of the particular program in question. My task as an evaluator, then, was not primarily aimed at producing information which is generalizable to other sorts of instructional programs. Thus, as long as my evaluative intention remains primarily improvement-oriented, information about other "universes" is relatively unimportant and the limitation on generalizability to such other "universes" will be of less concern.

Since it treats each instructional program as a unique entity, the socio-anthropological approach would initially appear to produce totally non-generalizable findings. If a secondary purpose of the evaluation study is generalization to other programs or settings, this may be a serious charge. It is countered with the argument that the uniqueness of a particular learning milieu does not preclude commonality of certain characteristics with other milieux. Even though the implementation of each instructional program is a unique entity, never to be repeated in exactly that form, it may share instructional methods, elements of content, instructors, physical circumstances, and so on, with a variety of similar programs. In the extreme, though they are somewhat different, the two workshops represented by Studies I and II were very similar. A significant amount of extrapolation or projection from one to the other, and generalization about other instructional programs which share similar characteristics, would appear to be reasonable. The descriptive nature of information produced by the socio-anthropological approach greatly assists in determining the extent of commonality between instructional sequences and between an instructional program and a hypothetical "universe of information". Thus, descriptive data assists in assessing the validity of any generalizations made. If the detailed descriptive data was not available we would be left with very little information about characteristics of the milieu upon which to assess the validity of the generalizations which might be

made. Since it does not focus attention on only a few elements within the instructional system, the paradigm also significantly increases the range of possibly valid generalizations which can be made. It encourages the investigator to procure information about a broad range of elements and issues, thus increasing the latitude of the information sample in validly representing other universes of information.

Earlier it was suggested that the reliability of information produced by the socio-anthropological approach is enhanced by its advocacy of "triangulation". If a number of different data-generating media and sources agree as to the dimensions of a phenomenon, the information set as a whole would appear to be relatively reliable--at least consistent. We have also noted that unreliable information does not incline toward valid generalization. To the extent that the socio-anthropological approach produces reliable information, it appears to increase the opportunity for valid generalization. On the other hand, to the extent that terms are not stringently defined in a socio-anthropological study, the precision of generalization is diminshed. If we do not specify very precisely what it is we have assessed, how we have assessed it, and what we found, our generalizations from those findings will necessarily reflect that imprecision. The reliance on subjectivity and tolerance of imprecise definition in the paradigm necessarily vitiates the exactitude of generalizations which flow from its use.

In summary, it would appear that, though use of the paradigm does impose some special limitations on the generalizability of data rendered, the scope and volume of that information provide opportunities for valid generalization which could not possibly exist in data obtained from a more circumscribed investigation.

Utility

If sufficient validity, reliability and generalizability can be demonstrated or are assumed to inhere in information produced by a well-conducted socio-anthropological study, still other considerations inhibit or commend utilization of the paradigm. The paradigm permits collection of extremely detailed descriptive information on program elements and processes, thus affording decision-makers the opportunity to conduct a detailed review of the instructional sequence at almost any convenient or necessary time. Further, it allows for intensive study of diverse program elements themselves. Rather than having to base decisions simply on validation data, as would be the case with information derived from a classico-experimental investigation, decision-makers may thus "dissect" the program and examine the quality of its elements. In order for such detailed examination to take place a relatively large accumulation of detailed information must be available and the socio-anthropological approach allows--some would even say "encourages"-- the collection of appropriately large amounts of detailed information. Furthermore, the rather broad

focus of the paradigm permits gathering of information on a very wide front. Evaluation studies need not concentrate attention on a few highly specified learning objectives and associated outcomes. Using the socio-anthropological approach, it is entirely possible to produce useful information on a large number of: interpersonal and inter-group interactions occurring in the milieu, important issues in the content, problems in application of instructional methods, strengths and weaknesses of material and/or instructors, and so on. Thus, utilization of the paradigm allows us to find answers for many questions which the classico-experimental approach ignored. In Study II, for example, the report: permits a rather detailed examination of the interpersonal interactions which occurred in Group B on Thursday afternoon, it provides information about the later reactions of Group B members to that session, it reports on Group B's presentation to the Friday morning plenary session, and it attempts to assess the relevant opinions and observations of tutors and non-Group B workshop members. In effect, it allows for assessments with regard to the claim that Group B members would benefit more from remaining together than they would from joining in with the general group reconstitution. At the same time, the information contained in the report permits assessments to be made with regard to a wide variety of other concerns: the amount of didactic input, the utility of heterogeneous working groups, the benefits to be had from plenary sessions, the appropriateness of the tutors' evaluation session, the

usefulness of the written instructional material, and so on. The necessity of qualitative data to meet decision-making needs in areas such as these is documented by even the most vociferous proponents of quantification. Campbell and Borush, for instance, note:

> The qualitative project history becomes absolutely essential to proper interpretation of research results. The common-sense knowing involved at each stage of converting human experience and action into numbers . . . becomes focal. The subjective impressions of the many participants and observers of the social experiment acquire a relevance equal to that of the computer output. . . . Thus, we join Scriven in the heretical belief that there are other ways of knowing than through experiment. . . . In many areas, experimental data are much less equivocal than common-sense knowing or passive observational statistics. . ..5

As suggested early in this Chapter, along with breadth, volume, and detail in information produced, the paradigm permits a degree of flexibility in operations far beyond that of the classico-experimental approach. Without this flexibility the Group B phenomenon in Study II, for example, might have received only the barest attention. The paradigm is apparently well suited to examination of unanticipated events and outcomes--a feat which no design resting on pre-specified objectives and instruments could achieve. It is highly unlikely that prior to the workshop one would have predicted emergence of the group reconstitution issue or the value found in participant-generated case material. Since it is not geared to coping with them, a pre-specified design could only ignore such phenomena; the flexibility of the socio-anthropological approach allowed us to concentrate as much attention upon them as their

emerging importance warranted.

This flexibility in operations also permits us to substitute one data-gathering medium for another when circumstances so dictate. Rather than ignore important information because our investigative method will not practically or possibly permit us to examine a subject as we wish, another more practicable method of investigation may be employed in that particular subject area. Again, the paradigm encourages use of a variety of information-gathering media to "triangulate" on a subject. If one medium does not, for one reason or another, adequately explore the area in question, one or more of the other media may make up the deficiency, or at least remedy a substantial segment of it. In Study II, for instance, I found it physically impossible to directly observe the workings of all four small groups simultaneously. Within the prescripts of the socio-anthropological approach it was, however, possible to rectify that otherwise unavoidable deficiency by collecting information about the unobserved groups through informal conversations with tutors and participants, and through interviews. One group session was, in fact, recorded on audio-tape because direct observation, though impossible, was highly desirable. One can forsee circumstances in which audio- and video-tapes, closed circuit television, or films could substitute for direct observation and might even be included in evaluation reports to provide decision-makers with relatively direct access to extremely detailed information on program phenomena. Such techniques

might also serve to minimize certain investigator-linked influences. The major point, however,is that an evaluation paradigm which severely limits the media available for information generation consequently limits the information available to decision-makers. The flexibility of the socio-anthropological approach counteracts this limiting tendency and provides decision-makers with information from a much wider assortment of media. This is not to suggest that decision-making will be rendered easier--the greater variety of information available is likely to include more contraindications and qualifications on any decision alternative--but it does permit decision-making to be significantly more well-grounded.

Unfortunately these advantages are not without their blemishes. The generation of massive amounts of detailed information can be wasteful of precious resources. Much time and energy may be used in producing almost duplicate sets of data which are of little additional value, except in assessing information quality. Furthermore, a large volume of descriptive information is unweildy, making it rather difficult to store and retrieve at will. It is unlikely that the whole of a large-scale program could practicably be examined utilizing the variety of media which were used in Study II. The drain on resources would probably be prohibitive and the data-overload problem could easily become entirely unmanageable. The bulkiness of the report on Study II indicates how serious such a problem might become. Large amounts of information must be recorded,

stored, retrieved, examined, and edited. In order for a report--even on a relatively short and small instructional program-- to be concise enough for efficient consumption, a great deal of material has to be summarized or omitted. The editing job becomes formidible. This is likely to render the production task on a large-scale program evaluation infinitely more difficult and time-consuming. Yet for "triangulations" to be effective in examining alternative viewpoints, and available for assessing the reliability of data, much duplication of effort is inevitable. Thus, a design which depends upon variety in information-producing media requires a greater, or at least a more intensive, effort in recording, collating, and analyzing information; further, it creates additional problems in data storage and retrieval.

Regardless of how detailed or useful, descriptions of program elements and the relationships between them are not the same as statements about the consequences of instructional activity. Inferences about consequences are necessary to many evaluative purposes. If adequately applied, the stringent controls of the classico-experimental paradigm provide strength to such inferences. In employing the socio-anthropological paradigm, the same efforts to control "extraneous" influences are not present; thus inferences may be weaker. It would appear, then, that in circumstances where inferences about outcomes or impacts are highly prized and stringent experimental controls can be maintained, the classico-experimental approach appears to

offer greater possibilities for producing the information that is desired. By comparison, the qualitative information generated in an "illuminative" study is likely to incorporate less precise definitions and more questionable assertions; making for less well-substantiated and precisely stated inferences.

As in the last Chapter, we must also ask, "How 'expensive' was Study II, and does the level of expenditure indicate anything about the general utility of the socio-anthropological paradigm itself?" Excluding "writing-up" time, it is estimated that a minimum of 150 man/hours of work were expended in the production of Study II. Although this estimate includes a little time spent by participants while being interviewed and completing follow-up questionnaires, the great majority of this work fell to me. Furthermore, over one-third of the time spent was concentrated within the 100 hour period during which the workshop was in progress. I spent at least 60 of those 100 hours engaged in direct observation of sessions and gatherings, conducting interviews, conversations with participants and tutors, and recording. Some of this effort was, needless to say, rather enjoyable work--particularly the informal conversations over drinks in the bar. It was, nonetheless, intensely exhausting. In order to observe as much as possible of the workshop milieu, it was necessary for me to almost <u>constantly</u> attend to both formal and informal interaction. I was, therefore, usually one of the last to retire in the evening and one of the first to arrive for breakfast in the

morning. Additionally, some time had to be made available at the end of each day for recording; this made for very late nights. The intensity of this experience left me exhausted by the time the workshop was complete. If the program had continued for another week, I would have had great difficulty in maintaining the high level of involvement. It should, however, be noted that the residential nature of the instructional program both significantly increased the intensity of my efforts and permitted me to gather substantially more information from informal conversations than would have been possible in other circumstances. Had the program been organized on a non-residential basis, I probably would have expended less energy but collected less information from informal conversations.

Organizing, analyzing and editing the information was at least as time-consuming as the actual data collection. This does not include time spent in report writing, which is also not included in the Study I estimate. It is worth reporting, however, that because of the volume and unweildiness of the information, the report on Study II took significantly longer to write than did all of the writing for Study I; and Study I included a good deal of rating information which was not a consequence of the classico-experimental approach. The data collation and analysis procedures required approximately the same amounts of time and effort in both Studies, but would have been significantly reduced in Study I had I been more experienced in the manipulation of statistical data and the application of statistical tests.

In all, I expended substantially more time and effort in collating and analyzing data for Study II than I did for the performance test segment of Study I; but in Study I other people expended more time and effort on data production than they did in Study II. Thus, it would appear that the necessary workload entailed in using the socio-anthropological approach, while possibly less in overall terms, drew much more heavily on a single resource--the investigator. Even though it taps a wider array of information sources, with more variety in media, on a larger number of issues, utilization of this paradigm appears to ordain that a larger proportion of the work is done by the investigating agency itself. Furthermore, the use of a variety of information-gathering media requires a wide range of skill--both technical and social--from the investigator. Should the requisite combination of skills prove expensive to purchase, costs will rise significantly.

As regards actual financial outlay, Study II was significantly less expensive than the performance test portion of Study I. This was primarily because it required substantially smaller expenditures on instrument construction and application. Direct observation, interview, informal conversation and examination of documents required very little financial outlay; only that necessary to purchase a few audio-tapes, pens and paper. Only production of the follow-up questionnaire entailed much financial expense and, because of the significantly smaller amount of printing, its requirements did not begin to approach the expenditures

necessary in either segment of Study I. It would appear that the socio-anthropological paradigm's general flexibility can be used to render its applications significantly less expensive. As long as it is appropriate to the circumstances, ethically acceptable, and not applied in isolation, any data-gathering device might be used. Thus, less expensive media may be employed, rather than those which require very heavy expenditures of time, effort, or financial resources. However, in larger socio-anthropological studies, the higher costs of data storage and retrieval resulting from the volume and descriptive nature of the information obtained, would probably increase expenses considerably. Furthermore, even in Study II it is obvious that expenses would have risen considerably had I been unprepared to undertake the evaluative effort, thus requiring a purchase of appropriate services.

This last point should emphasize the fact that certain characteristics of the investigative agency assume a high level of importance in the socio-anthropological paradigm. To the extent that the classico-experimental approach achieves its goal of objectivity in measurement, the personal characteristics of the investigator are relatively unimportant. When the socio-anthropological approach is used, however, the personal characteristics of the investigator become crucial. He must be able to competently employ a wider variety of observational media. It is likely that he will require a great deal more knowledge of the substantive area involved. He must display a high level of skill in interpersonal communication, unquestionable

integrity, and considerable diplomatic aplomb. He must apperceptively temper his own subjectivity in reporting while relying heavily upon hunches and subjective impressions in data-gathering. He must be willing to assume a very large proportion of the overall workload. The quality of the information derived from a socio-anthropological investigation appears to rest heavily on these personal characteristics of the evaluator. Consequently, he must be chosen and/or prepared for the task with a great deal of care. If an "expert" is contracted to do the job, his services may be expensive.

The necessarily eclectic nature of a socio-anthropological approach is bound to occasion investigatory practices which sometimes appear to be wasteful of resources. One is hardly likely to explore uncharted territory without occasionally becoming lost. The blazing of necessarily circuitous trails through tortuous terrain demands heavy expenditures of energy and will occasionally appear to be pointless--especiallywhen no destination has been specified. It is clearly a task of exploration rather than measurement. As such, it requires that "dead-end leads" be uncovered; just as useful paths will also be laid down. The discovery of an apparently unbridgable chasm or insurmountable summit may be just as significant as the detection of an unrestricted way forward. The problematic situation is not amenable to solution until it has been located.

This topographical analogy should lead us to the relatively obvious conclusion that the socio-anthropological

paradigm logically presents itself for use when our investigatory purpose is exploratory and descriptive; not measuremental or, perhaps, validatory. This has been its proponents' major argument; that in circumstances where the terrain is unknown or inadequately charted, a different approach to investigation is in order. No matter where it leads, exploration is not wasteful if our purpose is essentially exploratory. As is the case with any set of action principles, however, an approach may be wasteful if it is applied for inappropriate purposes.

Many of the decisions which must be taken with regard to improving instructional sequences in social work clearly require information generated from broad explorations. Much of the terrain is uncharted. We know very litte, for example, about the effects on students of adopting a "unitary" approach to professional education and training. Our knowledge of the cognitive or affective processes engendered in students by combinations of fieldwork and classroom inquiry is extremely limited. The ways in which the interests, abilities, and deficiencies of students interact with those of instructors are poorly understood, if at all. These are but a few examples of areas in which we have pressing information needs; areas where our decision-making needs demand investigations of instructional sequences which offer considerably more than mere tests of instructional effectiveness assessed against a handful of rather limited criteria.

In order to facilitate the generation of information well-grounded in the existence of phemonena, rather than in pre-ordained theory, the socio-anthropological paradigm does not begin from a specification of expected outcomes. It thus permits an initially broad-ranging exploration which eventually becomes narrowed to those aspects of the instructional program's structure and process repeatedly revealed as significant. It seems less likely, then, to totally disregard phenomena which were not originally anticipated. Thus, it would appear to be well-suited to examining instructional systems about which highly inconclusive evidence and controvertible theory are available beforehand. A vast number of instructional sequences in social work--one might even say social work education and training in general--are characterized by inconclusive evidence and controvertible theory. Generally, then, though a search of the relevant literature leads to the conclusion that evaluators of instruction in social work have rarely used this more "engaged" paradigm, our investigation suggests that irrelevance or lack of utility to evaluative purposes should not be inferred. It has, in fact, been asserted that: ". . . in the competition for dominance in the intellectual spheres, it is the only new force worthy of attention."[6] Many would argue that this is an exaggeration borne of frustration with the more traditional approach. However, the very recognition that a classico-experimental approach often does elicit frustration suggests that an alternative may be of value. We have seen how the socio-

anthropological approach is concerned to explore, analyze, and describe the structures and processes of instructional sequences in situ, rather than concentrating wholly on an assessment of instructional effectiveness. Although this may be characterized as a "new" approach to the systematic evaluation of social work instruction, it is definitely not a model which is totally alien to the profession. Social workers have long been concerned with exploring, analyzing and describing the structures and processes involved in professional social work practice. Furthermore, the principles and methods associated with a socio-anthropological approach are not antithetical to social work principles and methods. Participant observation and interview have become standard exploratory media in social work practice. Value orientations associated with "heurism", "problem-centredness", and "beginning where the client is" are axiomatic to most modern social work practice and closely associated with a socio-anthropological approach to investigation. The desirability of methodological eclecticism and cross-disciplinary interchange are becoming increasingly strong in social work and are highly compatible with the socio-anthropological approach. Thus, one might expect that social work educators and trainers could adopt and adapt socio-anthropological principles and methods with relative facility and comfort. As this approach to evaluation is considered, toyed with, seriously implemented, and evaluated, those involved in social work instruction will undoubtedly discover further paradigmic merits and demerits, strengths and limitations.

CHAPTER VIII

TERMINAL OBSERVATIONS

It was noted in Chapter III that generalized theoretical constructs such as the socio-anthropological and classico-experimental paradigms must be tested and observed repeatedly in numerous forms of application before any really authoritative conclusions are possible. Our concern in this dissertation with the paradigms' uses in evaluation aimed at instructional improvement in social work must be seen as only a very limited contribution to the development of such conclusions. Our concerns have been restricted to only one substantive area and one major intention. Thus, we cannot expect that even a very considerable number of applications and analyses within these boundaries would substantially affect formal methodological theory. At the same time, enquiries into the paradigms' applications in a number of substantive regions may be expected to generate and validate formal methodological theory far beyond its present point. Comparative analyses of methodological applications within substantive areas such as ours have definite contributions to make beyond substantive boundaries. More important within our own bailiwick, social work education and training is faced with compelling and numerous desires, needs and demands for evaluative information which will contribute to program improvement. Cries for evaluation have been legion; responses almost negligible.

Indeed, the scarcity of published evaluations within the literature on social work education and training leaves us with little directly relevant material against which to judge the conclusions which can be drawn from this enquiry. Since the primary intended audience for most evaluations is rather limited, even if a great deal of evaluation activity is carried out it is unlikely that much of it will be reported in the professional literature. This would appear to be of particular concern with regard to applications of the socio-anthropological approach. Such applications currently appear to be rarities within a sparsely populated field. In considering the implications of my investigation, then, I have had very little assistance from the literature on education and training in social work. There do, however, appear to be quite a number of implications which warrant further examination. So, this final Chapter must be approached as an enquiry itself. Far from presenting a set of immutable conclusions and prescripts, it can but explore some of the conclusions and implications of our analysis, relate them to apparently germane controversies and major concerns within social work education and training, and suggest areas within which further hypotheses should be developed; hypotheses which will require considerable further investigation before conclusions of real substance can be synthesized.

Obviously, there are a very large number of issues associated with systematic attempts to evaluate instruction in social work; many of which are not necessarily intertwined

with the use of either paradigm examined here. For our purposes, such issues must be considered tangential, and must be substantially ignored if we are to concentrate on differential paradigmic contributions to instructional improvement activity. Our primary concern must remain with the adopted subject matter: the two paradigms as they may be applied to evaluation aimed at the improvement of social work instruction.

As Bennett and Lumsdaine point out, "Objectively or subjectively, consciously or subconsciously, we must evaluate to operate."[1] The investigative approaches which have been labeled "classico-experimental" and "socio-anthropological" are both conceptual frameworks which may be used to guide consciously systematic efforts to generate information of value to decision-makers. Similarly, this dissertation represents an effort to generate information of value to those who must make decisions with regard to the selection and application of paradigms for the evaluation of social work instruction. When undertaking a comparative analysis of two such models, it is rather easy to fall into the trap of thinking in terms of "right" and "wrong". The classico-experimental approach may appear to be "right" in some respects but "wrong" in others. A decision to use the socio-anthropoloigical paradigm in a particular circumstance may be viewed as either "right" or "wrong". In attempting to contrast the two paradigms, it is easy to dichotomize and to overstate a case. In the practical life of evaluative activity, however, there are no real right or wrong answers.

Evaluation is much more complex than simply avoiding certain errors or biases; evaluation practice is composed of a series of decisions taken in order to achieve more of one desired effect than another, and more desirable effects than undesirable effects. In short, evaluation research must be a series of trade-offs. Viewing evaluation in this light, it is suggested that this final portion of our enquiry will prove more useful if we reject the "right or wrong" mentality, in favour of an exploration which attempts to point out areas of possible gain and other areas of possible loss which might accrue from adopting or rejecting either of the paradigms--or elements of the paradigms. Further, we can look at a few indications of worthwhile hypotheses about appropriate and less appropriate paradigmic applications. Thus, it may be possible to outline some tentative hypotheses with regard to more closely matching methodological resources to information needs and instructional circumstances.

Toward Guidelines for Evaluation

As the paradigms-in-application were contrasted, certain conclusions or implications which are highly suggestive of action principles emerged. While it is recognized that further exploration, validation and debate are required before detailed guiding principles can be enunciated, at least a simple itemization of points does appear to be in order.

- When the classico-experimental approach is adopted for evaluating a sequence of instruction, the objectives of various factions should not be overlooked. Instructor- and administrator-held goals need not be the only source of objectives. Students, social work clients, employers, professional groups, and so on are vitally involved, and their particular aims and interests may well be of relevance to informing the evaluator's task. Even an evaluation which is restricted to assessments of student learning cannot afford to ignore the importance of these other objective sources. Further, it is worth keeping in mind that the process of specifying objectives tends to narrow the comprehensiveness of the objective set quite radically. Perhaps more consideration should be given to means for assessing progress toward goals which do not involve operationalization and measurement.

- A classico-experimental design is likely to be of particular value in cases where the instructional sequence under study is viewed as an entity in and of itself, ignoring the complex interaction of its constituent elements; or when it is possible to manipulate only one "variable" in the instructional situation (e.g., length of sequence), holding other variables constant. Thus, it would appear to apply most usefully in evaluations with a "summative" focus; where the assessment of overall program effectiveness is of primary concern. Unless an integrated series of mini-investigations can be carried out, the paradigm does not appear to be a major resource for conducting comprehensive evaluations

with a more "formative" focus. We must also remain aware that the degree to which we can provide good measures does not necessarily relate to the importance of the objectives involved, and that sophisticated statistical methods will not necessarily result in more valid or understandable information. No amount of statistical manipulation will compensate for the selection of an inappropriate research design.

- Operational definitions of instructional objectives would appear to encourage greater reliability in measurement while, at the same time, they may reduce the validity of indicators in reflecting the intentions from whence objectives were derived. Broader statements of objectives are unlikely to offer prospects for the same reliability in measurement but will probably encompass a more appropriately wide range of instructional goals. In considering alternative approaches to evaluation research, criteria related to the comprehensiveness of desired information must be considered as well as rigour in measurement. In order to produce the desired degree of comprehensiveness in information, a classico-experimental approach may prove impractical since a very large number of operationally defined objectives--and consequently high expenses--would be necessary.

- If a classico-experimental approach is employed to assess cognitive or skills learning in social work, it should be insured that the tasks, and circumstances surrounding tasks, used to measure learning are reflective of the tasks

and conditions under which it is intended that new skills and knowledge will actually be put to use. There would appear to be little value in measuring learning by means which are not closely related to the ways in which such learning is intended to be used.

- The classico-experimental paradigm would appear to be of more use to those engaged in evaluating instructional sequences aimed at the learning of relatively discrete sets of practice skills (e.g., in-service training authorities) than to those involved with instruction which aims at developing a broader constellation of knowledge, affect and skill in the student (e.g., graduate schools of social work). In relatively limited "training" sorts of programs, measurable behavioral outcomes that adequately reflect the major intentions of instruction may be specified. The classico-experimental paradigm offers outstanding opportunities for rigourous validation studies on programs for which instructional objectives can be operationally defined.

- As an integral part of any evaluation adopting the classico-experimental approach, efforts should be made to facilitate the adequate collection of descriptive information about the major significant conditions under which the instructional program operates, as well as about the dimensions of the instructional sequence itself. As Reid and Shyne observe, "It does little good to state that treatment X had better results than treatment Y, if we cannot say of what the treatments consisted."[2] They might

have added that information about the context within which "treatment" took place is also vital.

- If adequate preparation is done to insure that the basic requirements of a classico-experimental study exist before planning is complete and the instructional program begins, it is much more likely that adequate controls can be maintained and that the research will be significantly less expensive. When such research is extrinsic to a fully planned or already operating instructional sequence, there are likely to be many more problems in creating and using a control group, enlisting participant co-operation, avoiding disruption of the instructional sequence, or even operationally defining objectives and finding appropriate measuring devices. Such difficulties may prove to be very expensive, if not impossible, to ameliorate and may eventually lead to abandonment of the research--or to the generation of virtually useless information.

- In using a control group, it is vitally important to determine not only how the "outcomes" of control subjects compare with those of subjects exposed to the instructional sequence under study, but also what significant events occurred within the control group sample during the instructional period. Control group subjects may have undergone other experiences which significantly influence the "outcomes" displayed. Further, analyses should be carried out to determine whether subgroups within the two samples were affected differentially.

- Where the effects of an instructional program are expected to be very small indeed, or where they are extremely difficult to measure, major emphasis in research might be placed on simply determining whether or not any effects are present. The use of a control group may not be warranted in such circumstances. If, in fact, the use of a control group will require expenditures of resources at levels detrimental to the major emphasis, it should be dropped. Resources are often extremely limited and situations are likely to arise in which a quest for inferential strength through the use of a control sample will seriously impair efforts to detect subtle or elusive effects.

- The evaluator who intends to present information in quantitative form should keep it in mind that consumers of evaluation studies often know little, if anything, about statistical tests and concepts. The investigator's responsibility to communicate results fully, accurately and with minimal misunderstanding is, therefore, an onerous one. Efforts should be made to report findings in terms which do not require technical sophistication for understanding.[3]

- An investigation which fails to discover any effects of an instructional program may be just as important as one which discovers a number of significant effects. Results of an adequately conducted study which indicate "no effect" or "zero effectiveness" may be extremely important in a situation where certain effects or levels of effectiveness were expected. On the other hand, it is entirely possible that an investigation will yield statistically

significant findings that are educationally insignificant.

- Particularly when undertaking a formatively oriented evaluation, the investigator should not feel obliged to accept instructors' statements of intent and program philosophy without question. The possibly different viewpoint which he contributes may considerably benefit both program and evaluation planning. Even in cases of more summatively oriented enquiries, his questioning of basic philosophical constructs and instructional goals may be of real assistance in sharpening definitions of program objectives.

- The evaluator adopting a classico-experimental approach should seriously consider possibilities for using data-gathering media and sources not traditionally associated with experimentation. His study may be supplemented immeasurably by the use of tactics like: student, colleague and administrator ratings; simulation technology; systems analysis; participant observation; surveys; and interviews. Utilization of the methods and concepts employed by, or implied in, such approaches may, in fact, assist in mitigating some of the classico-experimental paradigm's limitations and disadvantages.

- Although the need for qualitative information will arise in all studies, one should not denigrate attempts to generate quantitative evidence. "Hard" data which pertains to the magnitude of program impact may be a major resource for those who must make decisions with regard to program continuation, curtailment, expansion, or broad

modification.[4] Furthermore, though objectivity may be seen as an ideal never to be completely attained, there is value to be had from efforts to approximate that ideal. No endorsement of idealism as a guiding philosophy in research is intended, but the development of pragmatic operational definitions and the discovery of empirically-ascertainable indicators of effectiveness will help to ameliorate certain of our pressing information needs.

- Similarly, though the qualitative data generated by a socio-anthropological enquiry are likely to suggest new theoretical formulations and identify difficulties which might otherwise have gone undetected, they tend to provide relatively weak support for firm generalizations. Efforts should be made to supplement such information with statistical data which will lend greater precision. While "theoretical" or "focused" sampling has a definite and valuable purpose, the appropriate use of adequate statistical samples offers a step toward more reliable inferences.

- The essence of the socio-anthropological approach appears to be engagement with program participants as other human beings. For this reason, fieldwork in anthropology does not deal with "subjects".[5] In order to maximize the value of his own informed human judgement and that of his informants, the investigator must refrain from slipping into a role which inhibits authentic human engagement with participants. This in no way suggests that he must be involved in the same ways as other participants are with the processes and structures of the program. His role is that

of investigator, not program member. As experience in Study 2 indicates, participation as an instructor or student may well divert entirely too much energy from the investigative task, rendering the evaluator inattentive to many of the most significant events and relationships occurring in the program. However, to maintain the trust and respect of his informants, as well as to thoroughly utilize his own judgemental faculties, the evaluator must engage participants as he finds them; as fellow human beings, not as loci of observations or objects of study.

- An encompassing initial exploratory phase in the enquiry helps to ensure that the emergent phenomena most significant to evaluative purposes are identified for further "progressive focusing" and that the information generated is grounded in aspects of actual program implementation. Without a broadly ranging exploration, it is doubtful that unanticipated events or events of unanticiapted significance will become apparent. At the same time, the investigator's prior knowledge and theoretical orientation will necessarily play a part in determining what he attends to and judges to be of possible significance. The "illuminative" investigator does not begin tabla rasa. By conducting a panoramic initial exploration and appropriately matching his prior knowledge of hypotheses that could be studied with the actual phenomena found, researchable areas are identified. The value of early explorations is exemplified by Michael's experience in an illuminative study on social work education. She concluded that her ". . . early discoveries led to the

formulation of concepts which were crucial to the development of the research."[6]

- Investigations of the instructional processes themselves reveal their inherent complexity and underline the inappropriateness of the classico-experimental approach in generating much of the information which will be useful to decision-making concerned with alteration of program elements. Traditionally, the label attached to the "black box" of instructional treatment has been invoked to account for results. Illuminating the interior of that box suggests that it cannot be treated as a single unitary factor if efforts to improve it are to meet with any success.

- Particularly in highly dichotomized environments, the evaluator must strive to resist identification with one "side" or the other. In order to be accorded the necessary degree of trust and respect by participants with diverse allegiances, he must retain an essential neutrality; standing between the various factions. To accomplish this without appearing to be unsympathetic or theocratic can be a frustrating and problem-riddled undertaking. During the course of his involvement with participants he must struggle to understand diverse points of view while consciously abstaining from apparent or real co-option to any particular subgroup or factional value orientation. At the same time, he must remain integrally hominal; refusing to adopt an attitude which would make him appear to be somehow above mundane human concerns. Obviously, this can be a thin and precarious line to tread.

- In the reporting stage of a socio-anthropological investigation, it would appear to be very important that the investigator explicate the most significant contextual, structural and procedural phenomena which bear on the data-generation effort. These should include contextually-linked descriptions of his own behaviors and theoretical orientations. Exposition of the investigator's role, as perceived by himself and others, is essential to the reader's understanding of the dynamic complex from whence data was created. The most significant aspects of the evaluator's own experiential background, value orientation, areas of expertise, limitations, and theoretical framework are entirely germane considerations in understanding and judging the qualitative information produced. Critical research processes, theoretical principles, and methodological guidelines can be discussed in relation to the milieux in which they were applied. By reviewing the investigator's own part in the interaction as integral to the study, an added dimension is imparted to the data. Experience in Study II suggests, however, that in short, intensely involving instructional sequences it may be extremely difficult for the investigator to attend to and record the details of critical investigative activities as they are undertaken. When one in intently concentrating on observing and recording a myriad of fast-moving interactions between participants, awareness of self-imposed discipline and applications of substantially ingrained assumptions may not even emerge. Necessarily snap decisions about "leads" to be followed or

rejected may not receive much attention. Subtle effects linked to the presence of the investigator might easily go unnoticed. Here, the close similarities with difficulties in maintaining professional self-awareness in some areas of social work practice might be noted. Through conscious effort, consultation with "outsiders", and systematic periodic review, the investigator can but try to document his own activities and affective states so that the most apparently significant, at least, will be available for analysis and reporting. As in social work practice, conscious exercise of self-awareness faculties would appear to offer a key.

- During the course of a socio-anthropological investigation, open communication which forthrightly establishes the investigator's role would appear to aid, rather than hinder enquiries. Furthermore, it acts as a measure to prevent the occurrance of ethical difficulties associated with investigative duplicity, unanticipated harm to informants, and participant resentment. Michael, for instance, noted that it was the ". . . creation and maintenance of meaningful and mutually understood roles . . ." which structured and controlled the information she obtained.[7] Analysis suggested that her major "mistakes" had arisen when, from the participants' point of view, she had transgressed the boundaries of the role cast for her. In a somewhat similar undertaking, Kelly found that practical and ethical difficulties ensued when her perceived investigative role became clouded.[8] In Study II, I observed that workshop

members were least likely to confide in me at the times when my role was least well defined--very early in the workshop when my role had not been well established, and in the initial stages of "after hours" discussions when they were unsure as to whether my role had somehow changed. Once my unabashedly investigative role had been clearly established or re-established, participants were much more forthcoming and apparently comfortable in relating their observations and opinions.

- At the same time that the evaluator should be forthright about his role and authentic, he must be prepared to circumscribe his own expressions of opinion and observation. In a study of this sort, the investigator's role renders him privy to confidential information which, by its very nature, may compromisingly identify or otherwise harm certain individuals or groups if divulged. Clearly, the investigator is under moral restraint in such matters; but the practical implications of breaching informants' confidences should not be overlooked. If an investigator should even appear to betray a confidence, he may well find that the information available to his further enquiries becomes very rapidly limited. Furthermore, it is assumed that the investigator will be concerned to minimize his own effect on the milieu under study and maximize the purity of his informants' intimations. By refraining from expressions which betray his own opinions, knowledge and beliefs about the program, he is likely to less markedly influence the opinions and beliefs held by his informants. This may, however,

become an inordinately difficult task. Occasions are likely to arise where the investigator feels an extreme social or moral obligation to express himself. He is, after all, likely to be more knowledgable about the processes, structures and outcomes of the instructional sequence than anyone else. As such, his opinions or observations may be of special value. In other instances, participants may enquire after his opinions as a prelude to expressing their own. In each set of circumstances the investigator must carefully weigh the possible consequences of his expressions against purposes and needs in conducting the evaluation. There may be times, especially when evaluative purposes are primarily formative in nature, when frank revelations of observation and opinion will yield the greatest benefit. On other occasions, it is likely that the withholding of personal expressions, frustrating though that might be, will be the most salutary course of action. Experience in Study II indicates that participants quickly come to recognize and accept the necessity for such reticence if it is adequately explained from time to time.

- The validity and reliability of information obtained from informants, as well as that from the investigator-cum-observer, should be examined. The veracity of "second-hand" information is open to even more question than that of information obtained by direct observation. Much of the information generated in a socio-anthropological enquiry will have resulted from perceptions, interpretations and communications on the parts of informants before it was received, interpreted and communicated by the investigator. Thus, the

perceptive, interpretative and communicatory faculties of the informants, as well as those of the investigator--and dimensions of the communication processes themselves--are of interest. Enquiries into the informants' observational powers, interpretative proclivities, and communicatory competences might afford both investigator and decision-maker valuable evidence about information quality.

- Investigative techniques which involve the specification of material to which informants will respond may inadvertently introduce inappropriate structure or content--perhaps stemming from the technician's own value orientation or theoretical stance--to the information they generate. Efforts should be made to prevent the structure and content of observational tools, such as interview pro forma or questionnaires, from surreptitiously insinuating such tendencies into the data. If these tools are broadly and simply designed, having due regard for the shape and order exhibited by the exploratory data, it is likely that the information procured will contain very little instrument-induced bias.[9] It may, in fact, be wise to abstain, to the extent possible, from introducing stimuli of this sort until after the initial exploration phase has been completed.

- "Outside" researchers and consultants may provide a very useful service to the evaluator by questioning his methods, observations, and conclusions; and by serving as reference points against which orientations may be checked from time to time. In the case of relatively long-term studies, special value may be had in regular and ongoing

consultations with uninvolved colleagues; helping to maintain focus.[10] In a more short-term investigation like Study II, this may not be feasible during the data collection stage; however, such consultants might be used in later analysis and interpretation stages. In relation to Study II, for example, later discussions with my faculty supervisors and student colleagues suggested alternative interpretations of events and dimensions in the data which had not previously been considered.

- In using a variety of information sources and data-gathering media to "triangulate" on aspects of the instructional sequence and cross-check major findings, one should not expect that all data sources will agree with each other; although it would appear that the media are more likely to do so. The generation of one homogeneously integrated set of information is highly unlikely. Dissenting viewpoints and contraindications will probably appear. Furthermore, in searching for the major trends and tendencies indicated in the information set, one should not ignore or depreciate slices of data which do not fit the emerging patterns. They may offer glimpses of alternative and possibly valid observations and interpretations. Veracity is not the prerogative of the majority and, though it may seem unfashionably undemocratic, the examination of nonconforming intimations may prove to be exceedingly valuable in the long-run.

Prospects for Synergy

In the early stages of this investigation it was hoped that evidence might be uncovered to indicate that some

form of amalgamation or close collaboration between the two paradigms would be feasible. Throughout the enquiry possible points of conciliation, melding and "symbiotic" linkages have been sought. I have been forced to conclude, however, that while the two paradigms might be used to very great benefit in tandem, true syncretism is out of the question; certainly in the terms originally visualized. Differences between the evaluative purposes addressed by the two approaches, and their intrinsically different philsophical orientations appear to render them mutually exclusive on the theoretical level. The magnitude of these basic differences suggests that any attempt to combine them into a truly integrated model would result in disservice to one or both of them.

The two paradigms display their intrinsic incompatibility from the very outset of study. The classico-experimental approach begins with a specification of the areas which are to be investigated. It prescribes the specification and operational definition of program objectives which will serve to focus the entire investigation.

> The first stage in the evaluation of any course . . . is to set down what the teacher is trying to achieve--the aims and the objectives of instruction. These should be set down in clear, unambiguous terms describing the expected changes in behavior.11

Proponents of the socio-anthropological paradigm, on the other hand, take the position that objective definition of this sort is an overly restrictive and unrealistic approach to evaluation. On the first point, the classico-experimental paradigm apparently equates "evaluation" with "validation". Evaluation, however, is not concerned only with assessing

effectiveness; it is more broadly concerned with examining and describing the instructional sequence as it exists, as well as discovering its "effects", "effectiveness", and "efficiency".

> The aims of illuminative evaluation are to study the innovatory program: how it operates; how it is influenced by the various school situations in which it is applied; what those directly concerned regard as its advantages and disadvantages; and how students' intellectual tasks and academic experiences are most affected. It aims to discover and document what it is like to be participating in the scheme . . . and . . . to discern and discuss the innovation's most significant features, recurring concomitants and critical processes.12

The fact that the socio-anthropological paradigm concentrates attention on "innovatory" programs should not be construed to mean that it may be applied to "first-run" programs only. As we noted in Chapter III, by concentrating attention on the instructional program as it is implemented, and as necessarily set within a "learning milieu", the paradigm may be used to study virtually any instructional program. The major point, however, is that the socio-anthropological paradigm adopts a significantly wider view of evaluation; a less restricted definition.

On the second point, the classico-experimental approach is purported to be somewhat unrealistic because it assumes, often erroneously, an ability to define one's instructional goals clearly and to specify means by which their attainment may be measured. Trow and Litwak point out a number of existential circumstances which often invalidate such assumptions.[13] It is not suggested that the classico-experimental approach may not have a valuable use in our evaluative

armamentarium, but that various practical considerations--chiefly our lack of knowledge--tend to limit its serviceability.

Thus, while the classico-experimental evaluator views the specification of operationally defined instructional objectives as of fundamental importance, those adopting the socio-anthropological approach do not. They see it as an often restrictive and unrealistic injunction. No amount of pressure could force these two viewpoints into a unified equilateral framework.

There is a further dimension of the breach between the two paradigms which has been implied in this investigation but which requires further clarification because it appears to have very significant implications for our efforts to improve instructional systems. In order to understand the paradigmic incompatibility more clearly, a short examination of our own enquiry--a meta-examination, if you will, must be added.

This investigation has concentrated attention on the technologies which are commonly associated with the two paradigms. When the enquiry was initiated it was supposed that a focus on what was actually done--the "methods" used--would be of practical value both in developing some tentative guidelines for evaluative activity and as a vehicle for examining some of the issues involved in evaluative activity. It was quickly realized that the techniques examined were based on axioms. That is to say, it was noticeable that certain techniques and certain axioms appear to correlate

positively and that the axioms must be accepted if the techniques are to be applied appropriately. A part of my investigative purpose, then, was to use an examination of the techniques as a foil to shed some light on the axiomatic assumptions supporting them. When one more closely examines this method of enquiry, however, one realizes that the dimension of human values tends to be ignored. Although my enquiries with regard to evaluative purposes are based on a recognition that values are involved, my adoption of a "methodological" perspective--an orientation that implies that method can be understood in and of itself--sometimes led me to neglect the necessary influence of human involvement in creating and sustaining the techniques and concepts which comprise methodology. My methodological perspective was useful in that it allowed me to examine and reach useful conclusions about the techniques and concepts which might be used in efforts to evaluate social work instruction. However, perspectives have boundaries and this particular perspective tended to limit me from examining methods as sets of techniques and concepts which are products of human creativity and acceptance. The tendency was to regard methods as if they somehow existed independently from human values. This point may be somewhat difficult to see because obvious exceptions can be found in the investigation. In many other instances, however, the examination did tend to treat the methods involved as if they were actualities set apart from the context in which they necessarily existed; set apart from the involvement of people.

One of the clearest indications of this tendency is my temptation in reporting to say "the classico-experimental paradigm prescribes" or "the socio-anthropological paradigm suggests". Increasingly, it became apparent that it was the proponents of the classico-experimental paradigm who prescribed, the adherents to the socio-anthropological paradigm who suggested. It slowly became clear that the paradigms resulted entirely from the purposive action of human beings.

A consequence of my methdoloogical orientation was my delineation of "quality criteria" primarily in terms of effectiveness. The criteria chosen focused upon what the information or methods would do for us, not on what they should do for us and why. Quality was viewed in terms of "correctness" or "incorrectness". To a large extent, I had unwittingly adopted an ideological position which tends to ignore value judgements as a major element in the assignation of quality and ". . . involves discounting the possibility that what is 'rational, logical and practical' for one person may not be so conceived by another."[14] Had the content of my investigation been different, I might not have come to glimpse the limitations inherent in that approach.

The foregoing short examination of my investigation should lead us toward a deeper understanding of the basic breach between proponents of the classico-experimental paradigm and those who espouse the socio-anthropological paradigm. In adopting the classico-experimental approach to evaluation, events and relationships which occur during an instructional sequence, the "characteristics" of persons,

and the "outcomes" of instruction are often treated as if they were objects which somehow exist independent from human presence. In an effort to emulate the positivist approach of natural science, social and psychological phenomena may be treated as if they are qualitatively the same as objects. Human realities may be reduced to the order of things so that they might be studied by use of the same methodology that has proved to be so valuable in physics, chemistry, agriculture, botany, and so on. Trustees of the socio-anthropological orientation explicitly reject this "reification of constructs"; viewing human products, including instructional systems, as fully understandable only in relation to the social and psychological contexts in which they exist.[15] Thus, it would appear that the philosophical gap between adherents to the two paradigms is very wide indeed. Books and articles which document and explore that gap are appearing in ever-increasing abudance as dissatisfaction with a positivist approach to the study of human enterprise mounts.[16] Unfortunately, further exploration would rapidly move us away from the main thrust of this dissertation and into realms of philosophical argumentation which I am not prepared to review here.

In terms of synergy, it would be easy to conclude that the two paradigms do not represent compatible components for an encompassing, integrated theoretical construct; that each may represent a valid way of going about evaluation in certain circumstances, but that the purposes and defintions implied by their different terms of reference render the

models mutually exclusive. It should be noted, however, that at the operational level, combinations are definitely possible and probably desirable. Activity designed to produce appropriate qualitiative information to supplement and extend quantative data would appear to be entirely feasible, if slightly heretical, within the classico-experiemtnal tradition. The proper use of quantitative data to supplement and extend qualitative information is not only feasible but sometimes solicited by proponents of the socio-anthropological approach. Thus, combinations of the data-gathering techniques associated with each paradigm could be used to generate more complete evaluative information. If our purpose in evaluation was primarily conceived to be validatory, then techniques associated with the generation of qualitative information could be used to supplement a classico-experimental approach. If our purpose in evaluation was primarily conceived to require more than an assessment of effectiveness, then techniques which generate quantitative data could be used to supplement a socio-anthropological approach. A broadly conceived evaluation would probably use both sorts of data-generating technology.

Implications for Instructional Improvement

This investigation has adopted a primary concern with efforts to improve instruction in social work. Evaluation may provide information necessary to making decisions which affect improvements, but evaluation research does not do the whole job. Those decisions must be made; a planning process

must be undertaken. Basic difficulties arise not only in deciding what will constitute improvement, but also in deciding how the planning process should be structured. At the moment we would appear to have at least three workable alternative approaches to instructional planning. First, we can adopt a planning process based upon the epistemological orientation associated with the investigative paradigm labeled "classico-experimental". Second, we can utilize the framework which has been characterized as "socio-anthropological". Third, we can find ways of combining elements of both orientations to create a new approach to instructional planning. A fourth alternative may be available but does not appear workable within the framework adopted by this enquiry; that other perspectives on knowing, either currently available or yet to be developed, might be used to underpin the instructional planning process. Within the terms of reference adopted here, attention must be concentrated on the first three alternatives.

If one accepts the classico-experimental orientation as a framework within which to evaluate programs and conduct the planning of future instructional sequences, one adopts an essentially "rational" planning focus. One assumes a logical order of progression from problem identification, through selection of instructional goals and partialization of goals into specific instructional objectives that are related to instructional capacities, to implementation and evaluation. As Guzetta observes, this has been the traditional approach to curriculum planning in social work education.[17]

Such an approach endows our instructional planning efforts with a relatively well-defined structure and permits us to benefit from the strong points which were noted in the classico-experimental paradigm; most notably the inferential strength and qualities associated with easy data-manipulation.

Unfortunately, such an approach to the planning of social work instruction also suffers from the shortcomings observed in our examination of the classico-experimental paradigm in use. The principal drawback results from the circumstance that there is not as yet, a strong consensus within social work with regard to the particular human needs to be addressed, the interventive strategies to be used in meeting those needs, and the knowledge and skills most necessary to meeting such needs. As a result, the profession remains unclear about many of its desired goals and is often unable to formulate appropriately specific instructional objectives. Unless much greater clarity about instructional goals can be brought about--through investigation and conference on needs to be addressed, interventive modalities to be utilized, and professional knowledge, values and skills to be maximized--the "rational" approach to evaluation and curriculum planning will be of rather limited use. It was observed that lack of clarity about goals in Study I seriously restricted the usefulness of the classico-experimental paradigm's application. Precisely the same observation could be made about social work education and training in general. Within certain spheres (e.g., aspects of in-service training and specific skill development programs) goals and objectives may be spelled out

clearly. On the whole, however, instruction in social work lacks specific direction. We know that we want better social workers but we do not know what specifically makes for better social workers. Much more of our own research and much more conclusive debate would appear to constitute the way forward. Thus far, our limited research on practice effectiveness and our rather shallow explorations with regard to the implications of various value commitments have failed to provide the profession with a great deal of direction. If we wish to beneficially use a "rational" approach to instructional planning, a stronger sense of direction is imperative.

As opposed to the "safer" classico-experimental approach, a framework for instructional planning which incorporates the socio-anthropological perspective is likely to involve more risk-taking as well as more imaginative and visionary activity. As we noted in Study II, the paradigm permits greater flexibility and more rapid response to change. Thus, it would appear to offer the instructional planning process less structure but more capacity to adapt to changes in the society and in the profession. This increased flexibility would appear to be a very useful quality in a context where both professional and societal circumstances are changing rapidly. The socio-anthropological paradigm permits the evaluator to tailor his techniques very closely to the demands of the context in which he is working, and to respond rapidly to contextual change. In the same way, the instructional planner who adopts this more "heuristic" orientation will be permitted to fashion his program designs and revisions very

closely to the needs and demands of the profession and society; as well as to respond more quickly to environmental changes.

There are, of course, dangers in this approach as well. Rapid adaptation to professional and societal change may predispose the planner to creating instructional fads, corresponding to current societal or professional fads. In time, this may result in widespread disillusion with the planning process and greater confusion about directionality. The highly unstable curriculum is unlikely to encourage assiduous concentration on attaining long-range objectives. Furthermore, as may be adduced from my own frantic activity in Study II, such an approach may predispose the instructional planner to stretching his resources beyond bearable limits. Just as I toiled to almost utter exhaustion during Study II, the instructional planner may inadvertantly demand more effort or financial expenditure than the system can bear. If the demands are tempered to fit the availability of resources and if rewards for resource expenditure are perceived to be forthcoming, this may result in a maximal level of resource utilization. If that balance is not maintained, however, extreme dissatisfaction and disillusionment, again, may result. It would appear that wisdom gained through experience and careful detailed observation of the system in action are required to balance demand with resource availability.

As we noted earlier, a true amalgamation of the two paradigmic orientations does not appear to be feasible. However, it may be possible at a more technical level to use the approaches co-operatively; thus creating an overall

instructional planning process which exhibits the major strengths of both. If efforts are taken to avoid the pitfalls and difficulties inherent in each paradigm, perhaps a much more valuable planning process can be created. One way of attempting to do that has been suggested by Rothman and Vigilante.[18] They suggest an approach to curriculum development in social work which utilizes the "rational" and "heuristic" approaches in tandem. Very succinctly, their framework appears to involve supplementing activity based on a classico-experimental orientation with infusions of activity emanating from a socio-anthropological orientation. Careful examination and analysis of existing curricula is used to identify particular shortcomings and problems. This is followed by a systematic analysis of possible alternative routes for stregthening the curriculum, and a decision is taken to select one or more alternatives for experimentation prior to initiating a major curriculum change. It is assumed that the findings from these experiments will provide a valid base for undertaking final solutions.

> All of these steps would presumably take place in the order presented, all very rational. But in the meantime, a final decision dwells in limbo. Curriculum limbo is exemplified by unmet demand and discontent flooding the field, oftentimes reducing the experimentation undertaken for rational planning purposes to an inconsequential or obsolete level as the time lag between felt need and curriculum action widens. The rational approach often bogs down because the specimen won't stand still. The irrational consequences of rational planning seem to require the infusion of the heuristic dimension.19

Thus, Rothman and Vigilante's planner now turns to the "heuristic" mode and adopts what appears to be the most immediately feasible solution within the context which exists

at that time. Ongoing modifications to the instructional plan are adopted as illuminative investigations of continuing programs suggest redirections. This course of action continues until results of the experiments suggest a final solution. So, Rothman and Vigilante appear to see their proposal as delimiting an approach to curriculum planning strongly based in the classico-experimental mode but occasionally supplemented by use of the socio-anthropological orientation.

There are two assumptions and one omission in the foregoing proposal which lead me to believe that a somewhat different configuration makes more sense. The first assumption seems to be that social work education and training in general is not now going through a crisis. Evidence strongly suggests the opposite point of view. As we noted earlier, social work education appears to suffer from a lack of clear direction. There is a very great need for well-defined educational goals, and this is not being met adequately by current investigative and discursive activity. Our substantive knowledge base is inadequate, our "unique" skills ill-defined, and the implications of our value orientations inadequately explored. Simultaneously, society--primarily through governmental bodies - is increasingly demanding that we account for our collective behavior as a profession. Social work in general, and social work education in particular, is embroiled in a crisis of confidence and lack of direction.

The second questionable assumption in Rothman and Vigilante's model is that findings from the classico-experimental approach will provide a valid base for undertaking final solutions to instructional dilemmas. This may be entirely unwarranted because such an approach cannot contend with the element of contextual change. It tends to offer solutions to old problems, not current ones. As long as the environmental contexts within which social work instruction operates continue to change, and especially if that change is rapid, investigations in the classico-experimental mode will not provide final solutions. In fact, there never will be any final solutions.

The omission in Rothman and Vigilante's proposal is the lack of any suggestion as to how the initial examinations of existing curricula and possible alternative routes to strengthening curriculum are to be carried out. As we have seen, investigations do not proceed without reference to some overall framework, either implicit or explicit. The classico-experimental paradigm would appear to be rather poorly suited to these essentially exploratory tasks. Such efforts would almost undoubtedly bog down in difficulties arising from the scarcity of clearly specified instructional objectives in existing programs. Furthermore, that paradigm does not appear to offer a very profitable approach to identifying alternative courses of action. Although it may not offer such rigorous "control" and sophisticated statistical analysis, the socio-anthropological paradigm seems to represent a better approach to both tasks. It remains

useful in situations where instructional objectives are not clear, and it appears to be particularly useful for exploring alternative courses of action.

An alternative framework for instructional planning, then, might begin from the initial use of a socio-anthropological perspective to guide the exploration of existing instructional systems, seek means for improvement, and generate middle-range theory well-grounded in the context of social work education and training. Those resulting hypotheses that could be usefully honed down to include operationally defined instructional objectives might then be used in assessments of effectiveness, using the classico-experimental approach for both improvement and accountability purposes. Ongoing socio-anthropological investigations would be used to follow up leads contained in both sorts of findings and explore new dimensions and contexts. Thus, evaluations from both perspectives would gradually lead us toward a more integrated and comprehensive body of knowledge and theory about social work instruction. In this way, evaluation, planning and implementation might really be viewed as overlapping components in a continuum which gradually generates genuine improvements and helps to meet accountability demands.

Need for Further Study

Many areas requiring further study have been suggested by this investigation. On a broad front, investigation and further discussion in virtually every area concerning

social work instruction is desirable. We have noted that evaluation research is primarily aimed at the generation of useful information about particular programs or program elements, but the secondary concern of generating more broad-range knowledge and theory must not be ignored. Social work instruction does operate from a somewhat inadequate knowledge base, and its theoretical underpinnings display little integration. Evaluations of specific programs can play a part in developing a more adequate knowledge base and encouraging theoretical growth. On a more limited front, further evaluation and meta-evaluation can provide much-needed information which will assist in improving our evaluation methodology itself.

With regard to improving instruction in social work, much more attention must be paid to investigations of instructor and student activity in the instructional process. Our understanding of the cognitive, affective, and social activities associated with various forms of learning is wholly inadequate. The benefits, detriments, and problems involved in efforts to individualize learning have received little empirical attention. The differential utility of various instructional modes for teaching social work content (e.g., lectures, discussions, simulations, modeling) require a great deal of attention if instruction is to be highly effective and efficient. The usefulness of teaching abstract constructs to facilitate acquisition of practice skills has received inadequate attention. Explorations of current social work practice and more accurate prediction of future

practice needs will help us to more precisely define professional roles and delineate the skills necessary to successful role conduct; thus providing social work instruction with more specific directives for instructional planning and more specific objectives for evaluation.[20] The feasibility and desirability of integrating consumer input on all levels (e.g., student, client, agency, and profession) into instructional planning and evaluation will be of special importance--especially in meeting accountability demands.[21]

In terms of evaluation methodology, numerous issues must be addressed. Cost-effectiveness measures, participant resistance, internal versus external evaluation, standardization of tests, criterion- versus norm-referenced tests, and timing of evaluation onset are but a few of the areas requiring more exploration. Better means must be developed for assessing: self-directedness in learning, readiness to learn, mutualism between instructor and student, and a problem-centred orientation to learning. Social work instruction appears to be moving in these directions, but little attention has been paid to understanding either the means or implications that attend such movement. Finally, a much deeper understanding of approaches to evaluation is required. Our investigation of two paradigms strongly suggests that evaluation models must be carefully chosen for their appropriateness to the subject and purposes of evaluation, and also that they may require considerable adaptation when applied to any specific instructional program.

Thus, a detailed and well-sustantiated typology of evaluation methodologies--as applied to social work instruction--should be developed. The investigation has suggested quite a number of elements which might be used in such a typology. If evaluative dimensions like process/outcome orientation, formative/summative information needs, reliability/validity considerations, and anticipated costs of various sorts can be closely related with program variables such as length, type of objectives, instructional methods, and environmental conditions, the appropriate choice and adaptation of evaluative strategies for specific programs should become a better informed and more rational undertaking.

NOTES

NOTES

CHAPTER I

[1]Carol Weiss, Evaluation Research: Methods for Assessing Program Effectiveness (Englewood Cliffs, N.J.: Prentice-Hall, 1972), p. 11.

[2]Daniel L. Stufflebeam, "Alternative Approaches to Educational Evaluation: a Self-study Guide for Educators" in W. James Popham (ed.) Evaluation in Education: Current Applications (Berkeley: McCutchan, 1974), p. 121.

[3]Michael Scriven, "The Methodology of Evaluation" in Perspectives on Curriculum Evaluation, AERA Monograph Series on Curriculum Evaluation, No. 1 (Chicago: Rand McNally 1967).

[4]Stufflebeam, "Alternative", p. 122.

[5]See Scriven, "Methodology of Evaluation".

[6]This definition derives from the CIPP model presented by Daniel L. Stufflebeam, et al., Educational Evaluation and Decision-Making (Itaska, Ill.: F. E. Peacock, 1971). For a succinct exposition of the model, and exercises designed to promote understanding of it, see Stufflebeam, "Alternative", pp. 117-142.

[7]For further discussion, see Weiss, p. 5.

[8]For useful discussion see C. W. Churchman, The Systems Approach (New York, Delacorte Press, 1968); for discussion of direct applications to instruction see A. J. Romiszowski, The Systems Approach to Education and Training (Page: London, 1970).

[9]For example, see: A. H. Maslow and W. Zimmerman, "College Teaching Ability, Scholarly Activity and Personality", Journal of Educational Psychology, 47 (1956) pp. 185-189; D. Soloman, L. Rosenberg, and W. E. Benzdek, "Teacher Behavior and Student Learning", Journal of Educational Psychology, 55 (1964) pp. 23-30; M. H. Abt and G. L. Fahey, "A Measurement of College Instructor Behavior", Scientia Paedagogica Experimentalis, 5 (1968) pp. 32-39; A. T. Sharon, "Eliminating Bias from Student Ratings of College Instructors", Journal of Applied Psychology, 54 (1970) pp. 278-281; and C. W. Grant, "Faculty Allocation of Effort and Student Course Evaluations", Journal of Educational Research, 64 (1971) pp. 405-410.

[10]For example, see: P. L. Compeau, Selective Review of the Results of Research on the Use of Audiovisual Media to Teach Adults, (Strasbourg: Council of Europe, 1972); R. Dubin and T. C. Taveggia, The Teaching-Learning Paradox: A Comparative Analysis of College Teaching Methods, (Portland University of Oregon, 1968); C. H. Cherryholmes, "Some Current Research on Effectiveness of Educational Simulations", Review of Educational Research, 38 (1968) pp. 160-176; and R. Glaser and W. W. Coley, "Instrumentation for Teaching and Instructional Management", in R. M. W. Travers (ed.) Second Handbook of Research on Teaching (Chicago: Rand McNally, 1973) pp. 832-857.

[11]A range of dimensions is discussed in W. J. Gnagy, P. A. Chesebro, and J. A. Johnson, Learning Environments (New York: Rinehart and Winston, 1972).

[12]H. J. Wahlberg (ed.) Evaluating Educational Performance (Berkeley: McCutchan, 1974) analyzes a number of variables from the social environment. See also: J. Withall, "Development of a Technique for the Measurement of Socio-emotional Climate in Classrooms", Journal of Experimental Education, 17 (1949) pp. 347-361; and J. Whithall, "The Development of a Climate Index", Journal of Educational Research, 45 (1951) pp. 93-99. For a short discussion of physical factors, see: Robert Dreeben, "The School as a Workplace" in Travers, Second Handbook, pp. 450-473.

[13]See Kenneth O. Doyle Jr., Student Evaluation of Instruction (Lexington, Mass.: D. C. Heath & Co., 1975) p. 5.

[14]For example, see: J. P. Guilford, The Nature of Human Intelligence (New York: McGraw-Hill, 1967).

[15]B. S. Bloom (ed.), Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I, Cognitive Domain (London: Longmans, 1956); and D. R. Krathwohl, B. S. Bloom, and B. B. Masia, Taxonomy of Educational Objectives: The Classification of Educational Goals, Handook II, Affective Domain (U.S.A.: Longmans, 1964).

[16]Doyle, Student Evaluation, Table 1-1, p. 7.

[17]See discussions in: J. S. Bruner, R. R. Oliver, P. M. Greenfield, et al., Studies in Cognitive Growth (New York: John Wylie & Sons, 1966); B. S. Bloom, Taxonomy; and R. L. Ebel, Measuring Educational Achievement (Englewood Cliffs, N.J.: Prentice-Hall, 1965).

[18]For examples, see: E. R. Smith and R. W. Taylor, Appraising and Recording Student Progress, The Progressive

Education Association (New York: McGraw-Hill, 1942); and N. E. Gronlund, Measurement and Evaluation in Teaching (New York: McMillan, 1965).

[19]For more comprehensive discussions, see: G. G. Stern, "Measuring Noncognitive Variables in Research on Teaching" in N. L. Gage (ed.), Handbook of Research on Teaching (Chicago: Rand McNally, 1963) pp. 398-434; and S. B. Kahn and J. Weiss, "The Teaching of Affective Responses" in R. M. W. Travers, Second Handbook, pp. 759-804.

[20]See: W. A. Mehren and I. J. Lehmann, Standardized Tests in Education (New York: Holt, Rinehart & Winston, 1969); and Gilbert Sax, "The Use of Standardized Tests in Evaluation", in W. James Popham (ed.), Evaluation, pp. 245-308.

[21]For example, see: Rolf P. Lynton and Udai Pareek, Training for Development (Illinois: Irwin-Dorsey, 1967); Malcolm W. Warren, Training for Results (Reading, Mass.: Addison-Weseley, 1969); the "In-Basket" exercise is discussed by F. M. Lopez, Evaluating Executive Decision-Making (New York: American Management Association, 1966).

[22]No attempt at discussion will be presented here. See: H. Jeffreys, Scientific Inference (Cambridge University Press, second edition, 1937); H. Feigl and May Broadbeck (eds.), Readings in the Philosophy of Science (Appleton, Century-Crofts, 1953).

[23]For useful discussion, see: Herman Wold, "Causal Inference from Observational Data: a Review of Ends and Means" in M. C. Wittrock and David E. Wiley, The Evaluation of Instruction: Issues and Problems (New York: Holt Rinehart and Winston, 1970) pp. 351-390; and Hubert M. Blalock Jr., Causal Inferences in Nonexperimental Data (Chapel Hill: University of North Carolina Press, 1964).

[24]Stufflebeam, Educational Evaluation.

[25]B. S. Bloom, J. T. Hastings, and G. F. Madaus, Handbook on Formative and Summative Evaluation of Student Learning (New York: McGraw-Hill, 1971) p. 11.

[26]For example, see: E. F. Lindquist (ed.), Educational Measurement (Washington: American Council on Education. 1951); R. M. W. Travers, How to Make Achievement Tests (New York: Odyssey Press, 1950); J. Heywood and A. H. Illife, Some Problems of Testing Academic Performance (Department of Higher Education, University of Lancaster, 1966); D. A. Wood, Test Construction (Columbus, Ohio: Charles E. Merril, 1960); and L. E. Taylor, Tests and Measurements, second edition (Englewood Cliffs, N.J.: Prentice-Hall, 1971).

[27]See sources in Note 20; these sources cover a range of issues which should be reviewed before any application takes place.

[28]For further discussion, see: Robert L. Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education, third edition (New York: John Wylie & Sons, 1969); Krathwohl, Bloom and Masia, Taxonomy; and B. O. Smith, "Teaching and Testing Values" in Proceedings of the 1965 Invitational Conference on Testing Problems (Princeton, N.J.: Educational Testing Service, 1960), pp. 50-59.

[29]For reviews, see: E. A. Bilodeau, Acquisition of Skill (New York: Academic Press, 1966); and Michael Posner and S. W. Keele, "Skill Learning" in R. M. W. Travers, Second Handbook, pp. 805-831.

[30]See sources in Note 21. Very useful articles may also be found in Journals such as: Training and Development Journal; Canadian Training Methods; and Training Directors Journal.

[31]L. E. Taylor, Tests and Measurements, p. 85.

[32]Doyle, Student Evaluation, p. 20.

[33]For critical discussion, see: M. Rodin, and B. Rodin, "Student Evaluations of Teachers", Science, 177, 1972, pp. 1164-1166; and C. F. Levinthall, L. M. Lansky and O. E. Andrews, "Student Evaluations of Teacher Behaviors as Estimations of Real-ideal Discrepancies: a Critique of Teacher Rating Methods", Journal of Educational Psychology, 62, 1971, pp. 104-109.

[34]K. O. Doyle and S. E. Whitely, "Student Ratings as Criteria for Effective Teaching", American Educational Research Journal, 11 (3), 1974, pp. 259-274.

[35]Other very useful information on a variety of techniques, may be found in: P. C. Smith and J. C. Kendall, "Retranslation of Expectations: an Approach to the Construction of Unambiguous Anchoring of Rating Scales", Journal of Applied Psychology, 47, 1963, pp. 149-155 (scaled expectations); J. C. Nunnally, Psychometric Theory (New York: McGraw-Hill, 1967) (forced-choice scales); K. S. Crittendon and J. L. Norr, "Student Values and Teacher Evaluation", Sociometry, 36, 1974, pp. 143-151 (importance ratings); F. Gagne and D. Allaire, "Summary of Research Data on the Reliability and Validity of a Measure of Dissatisfaction Derived from Reality-Desires Discrepancies", Institute National de la Recherche Scientifique, Universite du Quebec, 1974 (double-scale format).

[36]Maslow and Zimmerman, College"; L. M. Aleamoni and M. Yimer, "An Investigation of the Relationship Between Colleague Ratings, Student Rating, Research Productivity, and Academic Rank in Rating Instructional Effectiveness", Journal of Educational Psychology, 64, 1973, pp. 274-277.

[37]A. T. Drucker and H. H. Remmers, "Do Alumni and Students Differ in Their Attitudes Toward Instructors?", Journal of Educational Psychology, 42, 1951, pp. 129-143; J. Centra, "College Teaching: Who Should Evaluate It?", Findings, 1 (1), Educational Testing Service, Princeton, N.J., 1974.

[38]Doyle, Student Evaluation, p. 73.

[39]A. W. Bendig, "Student Achievement in Introductory Psychology and Student Ratings of the Competence and Empathy of Their Instructors", Journal of Psychology, 36, 1953, pp. 427-433; G. D. Lovell and C. F. Haner, "Forced-choice Scales Applied to College Faculty Rating", Educational and Psychological Measurement, 15, 1955, pp. 291-304; N. L. Gage, "The Appraisal of College Teaching", Journal of Higher Education, 32, 1961, pp. 17-22; N. F. Rayder, "College Student Ratings of Instructors", Journal of Experimental Education, 37 (2), 1968, pp. 76-81; Doyle and Whitely, "Student Ratings".

[40]A. T. Sharon and C. J. Bartlett, "Effect of Instructional Conditions in Producing Leniency on Two Types of Rating Scales", Personnel Psychology, 22, 1969, pp. 251-263; and Sharon, "Eliminating Bias".

[41]B. Rosenshine and N. Furst, "The Use of Direct Observation to Study Teaching" in R. M. W. Travers, Second Handbook, pp. 122-183.

[42]D. M. Medley, and H. E. Mitzel, "Measuring Classroom Behavior by Systematic Observation" in Gage, Handbook, p. 248. For a summary of "reactive effects" see E. J. Webb, et al., Unobstrusive Measures: Nonreactive Research in the Social Sciences (Chicago: Rand-McNally, 1969).

[43]For further discussion, see: R. W. Heynes and R. Lippitt, "Systematic Observational Techniques", in G. Lindzey (ed.), Handbook of Social Psychology (Cambridge, Mass.: Addison-Wesley, 1954), pp. 370-404; and R. M. W. Travers, An Introduction to Educational Research, third edition (London: McMillan, 1969), Ch. 3.

[44]Malcolm Parlett and David Hamilton, "Evaluation as Illumination: a New Approach to the Study of Innovatory

Programs", Centre for Research in the Educational Sciences, University of Edinburgh, 1972, p. 23.

[45]For further discussions, see: C. H. Backstrom and G. D. Hursh, Survey Research (Evanston, Ill.: Northwestern University Press, 1963); A. N. Oppenheim, Questionnaire Design and Attitude Measurement (London: Heinemann 1966); and D. R. Berdie and J. F. Anderson, Questionnaires: Design and Use, (Metuchen, N. J.: Scarecrow Press, 1974).

[46]Exposition of this argument may be found in Doyle, Student Evaluation, p. 12.

[47]For a discussion of definitions, see Robert Ebel, "Must all Tests be Valid?", American Psychologist, 16, 1961, pp. 640-647.

[48]Clinton I. Chase, Measurement for Educational Evaluation (Reading, Mass: Addison-Wesley, 1974), p. 64. For additional discussion, see: J. C. Stanley and K. D. Hopkins, Educational and Psychological Measurement and Evaluation (Englewood Cliffs, N. J.: Prentice-Hall, 1972), Ch. 4.

[49]Doyle, Student Evaluation, pp. 49-53. For further discussion, see: G. C. Helmstadter, Principles of Psychological Measurement (New York: Appleton, Century-Crofts, 1964), Ch. 4.

[50]Lee J. Chronbach and Paul Meehl, "Construct Validity in Psychological Tests", Psychological Bulletin, 52, 1955, pp. 281-320; and Lee J. Chronbach, Essentials of Psychological Testing, third edition (New York: Harper & Row, 1970).

[51]Chase, Measurement, pp. 59-63, discusses the use of statistical correlations as "coefficients of validity"; while useful and thorough discussion is provided by Lee J. Chronback, "Validation of Educational Measurements" in Proceedings of the 1969 Invitational Conference on Testing Problems (Princeton, N.J.: Educational Testing Service, 1969), pp. 35-52.

[52]For discussions, see: R. L. Thorndike, "Reliability" in E. F. Lindquist, Educational Measurement, Ch. 15; F. B. Davis, Educational Measurements and Their Interpretation (Belmont, Calif.: Wadsworth, 1964), Chs, 2,4,10,13 and Appendix F.

[53]For definitive sources, see: G. F. Kuder and M. W. Richardson, "The Theory of the Estimation of Test

Reliability", Psychometrika, 2, 1937, pp. 151-160; and D. P. Hoyt, "Test Reliability Obtained by Analysis of Variance", Psychometrika, 6, 1941, pp. 153-160.

[54]Doyle, Student Evaluation, p. 34.

[55]J. P. Guilford, Psychometric Methods (New York: McGraw-Hill, 1954), pp. 280-288. See: two articles by I. I. Bejar and K. O. Doyle Jr., "The Effect of Prior Expectations, on the Structure and Evaluation of Student Ratings of Teacher Behavior" and "Expectations, First Impressions and Evaluations", Measurement Services Center, University of Minnesota, Minneapolis, 1974; R. F. Burnsaka and T. D. Hollmann, "An Empirical Comparison of the Relative Effects of Rater Response Biases on Three Rating Scale Formats", Journal of Applied Psychology, 59 (3), 1974, pp. 307-312; J. A. Centra, "Self Rating of College Teachers: a Comparison With Student Ratings", Journal of Educational Measurement, 10, 1973, pp. 287-295; R. K. Isaacson et al., "Dimensions of Student Evaluations of Teaching", Journal of Educational Psychology, 55 (6), 1964, pp. 344-351; Sharon "Eliminating Bias"; and R. D. Rees, "Dimensions of Students' Points of View in Rating College Teachers", Journal of Educational Psychology, 60, 1969, pp. 476-482.

[56]See any good textbook on social research: e.g., S. Labovitz and R. Hagedorn, Introduction to Social Research (New York: McGraw-Hill, 1971), Ch. 3.

[57]Weiss, p. 10.

[58]For a review and selected annotated bibliography, see Richard I. Miller, Evaluating Faculty Performance (London: Jossey-Bass, 1972).

[59]Doyle, Student Evaluation, p. 85.

CHAPTER II

[1]Bertha Caplan Reynolds, Learning and Teaching in the Practice of Social Work (New York: Farrar & Rinehart, 1942).

[2]Helen Harris Perlman, "Teaching Casework by the Discussion Method", Social Service Review, XXIV, (3), 1950, pp. 334-346.

[3]Helen Harris Perlman, "The Lecture as a Method in Teaching Casework", Social Service Review, XXV, (1), 1951, pp. 19-32.

[4]Council on Social Work Education, Curriculum Study, Vols. I-XII, New York, 1959; See also. Eileen A. Blackey, "The Use of Audio-Visual Aids in Training", Social Casework, XXI (9), 1950, pp. 366-371; K. M. Forbes, "Training and Function", British Journal of Psychiatric Social Work, 2 (5), 1951; Grace L. Coyle, "The Role of the Teacher in the Creation of an Integrated Curriculum", Social Work Journal, 33 (2), 1955, pp. 73-82; Alan F. Klein, Role Playing in Leadership Training and Problem Solving (New York: Association Press, 1956); Eileen A. Blackey, Group Leadership in Staff Training (Washington D.C.: Department of Health, Education and Welfare, Bureau of Family Services, 1956); Pauline D. Lide, "An Experiment in Teaching Social Casework", Social Service Review, XXXIII (3), 1959, pp. 253-259.

[5]For a full copy of Tyler's text, see: R. W. Tyler, "Distinctive Attributes of Education for the Professions", Social Work Journal, 33 (2), 1952, pp. 55-62. It is also reprinted in A Source Book of Readings on Teaching in Social Work (New York: Council on Social Work Education, 1966).

[6]Charlotte Towle, The Learner in Education for the Professions (Chicago: University of Chicago Press, 1954).

[7]Towle, The Learner, pp. 25-26.

[8]Alfred Kadushin, "Interview Observation as a Teaching Device", Social Casework, XXXVII (7), 1956, pp. 334-341.

[9]Martha Moscrop, In-Service Training for Social Agency Practice (Toronto: University of Toronto Press, 1958).

[10]Elizabeth Herzog, Some Guidelines for Evaluative Research: Assessing Psycho-Social Change in Individuals (Washington D.C.: Department of Health, Education, and Welfare, Children's Bureau, 1959).

[11]Herzog, p. 2.

[12]Herzog, p. 83.

[13]Daniel S. Sanders, "Educating Social Workers for the Role of Effective Change Agents in a Multicultural, Pleuralistic Society", Journal of Education for Social Work, 10 (2), Spring 1974, p. 88.

[14]As an example, see: E. Matilda Goldberg, Ann Mortimer and Brian T. Williams, Helping the Aged: A Field Experiment in Social Work (London: George Allen & Unwin, 1970). For reviews and discussions of various studies, see: Joel Fischer, The Effectiveness of Social Casework (Springfield, Ill.: Charles C. Thomas, 1976); Edward J. Mullen, "Evaluative Research on the Effects of Professional Intervention: in Paul C. Vrooman (ed.), Trans-Disciplinary Issues in Social Welfare (Waterloo, Canada: Graduate School of Social Work, Waterloo Lutheran University, 1972), pp. 17-35; A. N. Grey and H. E. Dermody, "Reports of Casework Failure", Social Casework, November 1972, pp. 534-543; Noel Timms, "Research Into Social Work: A review article", British Journal of Social Work, 1 (3), 1971, 345-351; H. S. Maas (ed.), Research in the Social Services: A Five Year Review (New York: National Association of Social Workers, 1971); H. S. Maas (ed.), Five Fields of Social Service: Review of Research (New York: National Association of Social Workers, 1966).

[15]Marguerite Pohek (ed.), The Teacher's Compendium (New York: Council on Social Work Education, 1963). For an excellent discussion of the dimensions and implications of an andragogical stance, see: Malcolm S. Knowles, "Innovations in Teaching Styles and Approaches Based Upon Adult Learning", Journal of Education for Social Work, 8 (2), Spring 1972, pp. 32-39.

[16]Eileen Blackey, "Selection and Preparation of Faculty for Schools of Social Work", Journal of Education for Social Work, 1 (1), Spring 1965, p. 9.

[17]Alfred Kadushin, "Two Problems of the Graduate Program: Level and Content", Journal of Education for Social Work, 1 (1), Spring 1965, pp. 33-46.

[18]Herman D. Stein, "Cross-Currents in Practice, Undergraduate and Graduate Education in Social Work", Journal of Education for Social Work, 1 (1), Spring 1965, pp. 56-67.

[19]Mary E. Burns, "Paths to Knowledge: Some Prospects and Problems", Journal of Education for Social Work, 1 (1), Spring 1965, pp. 13-17.

[20]Burns, p. 13.

[21]Samuel Finestone, "A Memorandum on Research Into Social Work Education", Journal of Education for Social Work, 1 (1), Spring 1965, pp. 18-25.

[22]Finestone, p. 23.

[23]Finestone, p. 23.

[24]Novia Carter and Brian Wharf, Evaluating Social Development Programs (Ottawa: Canadian Council on Social Development, 1973). See also: Peter H. Rossi, "Practice, Method and Theory in Evaluating Social Action Programs" in James L. Sundquist (ed.), On Fighting Poverty (New York: Basic Books, 1969); and P. H. Rossi and Walter Williams, Evaluating Social Programs (New York: Seminar Press, 1972).

[25]Margaret Heyman, Criteria and Guidelines for the Evaluation of In-Service Training (Washington D.C.: Department of Health, Education, and Welfare, 1967).

[26]Dorothy Pettes, Supervision in Social Work (London: George Allen & Unwin, 1967).

[27]Mary Louise Somers, "Contributions of Learning and Teaching Theories to the Explication of the Role of the Teacher in Social Work Education", Journal of Education for Social Work, 5 (2), Fall 1969, pp. 61-73.

[28]Somers, p. 70.

[29]Marguerite Pohek (ed.), Teaching and Learning in Social Work Education (New York: Council on Social Work Education, 1970).

[30]N. L. Gage, "Successful Teacher Behavior: Can Science Contribute to the Art of Teaching" in Pohek, Teaching and Learning, pp. 1-12.

[31]N. L. Gage, "Successful Teacher Behavior", p. 10.

[32]Florence Clemenger, "Assessment of Students' Performance in the Classroom" in Pohek, Teaching and Learning, p. 33.

[33]For a notable exception, see the discussion in Peter W. Chommie and Joe Hudson, "Evaluation of Outcome and Process", Social Work, November 1974, pp. 682-687.

[34]Ronald Feldman, "Toward the Evaluation of Teaching Competence in Social Work", Journal of Education for Social Work, 8 (2), Spring 1972, pp. 5-15.

[35]Feldman, pp. 7-8.

[36]Feldman, p. 14.

[37]Charles S. Levy, "A Framework for Planning and Evaluating Social Work Education", Journal of Education for Social Work, 8 (2), Spring 1972, pp. 40-47.

[38]Levy, p. 41.

[39]Levy, p. 46.

[40]Walter L. Walker, "Changing Thought and Action Styles of Students and Faculty: Imperatives for Social Work Education?", Journal of Education for Social Work, 8 (2), Fall 1972, p. 57.

[41]Donald P. Hoyt, The Relationship Between College Grades and Adult Achievement (Iowa City, Iowa: American College Testing Program Research Report No. 7, 1965).

[42]Catherine Papell, "Teaching the Technology of Social Work: Cumulative Methods to Unitary Whole" in Kay Dea (ed.), New Ways of Teaching Social Work Practice (New York: Council on Social Work Education, 1972), pp. 1-18.

[43]Lola Selby, "The University of California Generalist Model for Teaching Social Work Methods" in Kay Dea (ed.), New Ways, pp. 19-32.

[44]Jack Stumpf, "Developing the Social Work Practice Curriculum at San Diego State College: A Further Statement", in Kay Dea (ed.), New Ways, pp. 33-47.

[45]Jack Stumpf, p. 47.

[46]The reader who wishes to further pursue this literature will find value in: L. Glick (ed.), Undergraduate Social Work Education for Practice: A Report on Curriculum Content and Issues, Syracuse University School of Social Work (Washington D.C.: U.S. Government Printing Office, 1971); Alvan M. Block, "The Dilemma of Social Work Education: Restructuring the Curriculum", Journal of Education for Social Work, 8 (1), Winter 1972, pp. 18-23; Arnold Gurin, "Education for Changing Practice" in Alfred J. Kadushin (ed.), Shaping

the New Social Work (New York: Columbia University Press, 1973) pp. 169-198; Phillip A. Fellin, "Long-Range Planning Implications for Social Work Education", Journal of Education for Social Work, 10 (2), Spring 1974, pp. 27-33; B. Rothman and J. L. Vigilante, "Curriculum Planning in Social Work Education", Journal of Education for Social Work, 10 (2), Spring 1974, pp. 76-85; Gwen Andrew, "Forecasting Social Work Practice as a Base for Curriculum Development, Journal of Education for Social Work 10 (3), Fall 1974, pp. 3-8; Ann McLean and Joan Robertson, "Systematic Integration of Consumer Input into Curriculum Revision", Journal of Education for Social Work,11 (2), Spring 1975, pp. 84-90; John Lewis and Faith Gibson, "The Teaching of Some Social Work Skills: Toward a Skills Laboratory", British Journal of Social Work, 7 (2), Summer 1977, pp. 189-209.

[47]Boyd E. Oviatt (ed.), Evaluation and Accountability in Social Work Education (Salt Lake City: Graduate School of Social Work, University of Utah, 1973).

[48]Gordon Hearn, "Accountability in Social Work Education" in Oviatt (ed.), pp. 4-15.

[49]Gordon Hearn, p. 5.

[50]Gordon Hearn, p. 9.

[51]Gordon Hearn, p. 15.

[52]Richard P. Lindsay, "Comments" in Oviatt (ed.), pp. 15-21.

[53]Richard Lodge, "Towards Accountability and Evaluation in Social Work Education" in Oviatt (ed.), pp. 23-26.

[54]Richard Lodge, p. 26.

[55]Rex A. Skidmore, "Evaluation and Accountability in Schools of Social Work", in Oviatt (ed.), pp. 27-34.

[56]Horace Lundberg, "Comments" in Oviatt (ed.), pp. 34-38.

[57]Horace Lundberg, p. 37.

[58]M. W. Linn, S. Stein, J. S. Carter, and B. S. Linn, "The C.A.R.E. Project: Interdisciplinary and Community-Oriented Approach to Training Social Work Students", Journal of Education for Social Work, 19 (2), Spring 1974, pp. 47-50.

[59]For discussion of these scales, see: C. E. Osgood, G. J. Suci, and P. H. Tannenbaum, The Measurement of Meaning (Urbana, Ill.: University of Illinois, 1957); and N. Kogan, "Attitudes Toward Old People: the Development of a Scale and an Examination of Correlates", Journal of Abnormal Psychology, 62, 1961, pp. 616-622.

[60]Michael F. Kolevzon, "Integrational Teaching Modalities in Social Work Education: Promise or Pretence?", Journal of Education for Social Work, 11 (2), Spring 1975, pp. 60-67.

[61]Kolevzon, pp. 62-63.

[62]George S. Sharwell, "Can Values be Taught? A Study of Two Variables Related to Orientation of Social Work Graduate Students Toward Public Dependency", Journal of Education for Social Work, 19 (2), Spring 1974, pp. 99-105.

[63]See: Barbara K. Varley, "Socialization in Social Work Education", Social Work, 8 (3), July 1963, pp. 102-109; Malcolm J. Brown, "Social Work Values in a Developing Country", Social Work, 15 (1), January 1970, pp. 107-112; Dorothy D. Hayes and Barbara K. Varley, "The Impact of Social Work Education on Students' Values", Social Work, 10 (3), July 1965, pp. 40-46; and Barbara K. Varley, "Social Work Values: Changes in Value Commitments of Students from Admission to M.S.W. Graduation", Journal of Education for Social Work, 4 (2), Fall 1968, pp. 67-76.

[64]Esther Sales and Elizabeth Navarre, "Teaching in Groups, Individual and Group Supervision in Field Instruction: A Research Report", School of Social Work, University of Michigan, 1970.

[65]Harris K. Goldstein and Linda Horder, "Suggested Teaching Plans for Maximizing Research Learning of Three Types of Social Work Students", Journal of Social Work Education, 10 (3), Fall 1974, pp. 30-35.

[66]See Harris K. Goldstein, "The Measurement of Attitudes and Research Knowledge", Tulane Studies in Social Welfare, New Orleans, Tulane University, 1968.

[67]Goldstein and Horder, p. 30.

[68]Raymond J. Thomlison and Frederick W. Seidel, "An Experiment in Curriculum Innovation in Graduate Social Work Education", Journal of Education for Social Work, 10 (3), Fall 1974, pp. 93-98.

[69]Thomlison and Seidl, p. 96.

[70]Alfred Kadushin and George Kelling, "Final Report, An Innovative Program in Social Work Education: The 3-2 Program", Madison, University of Wisconsin, mimeographed. For a summarization, see: Alfred Kadushin and George Kelling, "An Innovative Program Reducing Length of Training: Evaluation Procedures and Outcome", Journal of Education for Social Work, 13 (2), Spring 1977, pp. 68-75.

[71]Kadushin and Kelling, "An Innovative Program", p. 69.

[72]J. Brian Tully, "Personal Construct Theory and Psychological Changes Related to Social Work Training", British Journal of Social Work, 6 (4), 1976, pp. 481-499.

[73]See: M. Lifshitz, "Quality Professionals: Does Training Make a Difference? A Personal Construct Theory Study of the Issue", British Journal of Clinical Psychology, 13, 1974, pp. 183-189; A. Ryle and D. Breen, "Change in the Course of Social Work Training: A Repertory Grid Study", British Journal of Medical Psychology, 47, 1974, pp. 139-148; A. Ryle and D. Breen, "Social Work Tutors' Judgement of Their Students", British Journal of Medical Psychology, 47, 1974, pp. 149-152.

[74]John Carrier and Ian Kendall, "Social Policy and Social Change--Explanations of the Development of Social Policy", Journal of Social Policy, 2 (3), 1973, pp. 209-224.

[75]Carrier and Kendall, p. 222. For a much more complete discussion of this approach, see: Barney G. Glaser and Anselm L. Strauss, The Discovery of Grounded Theory: Strategies for Qualitative Research (Chicago: Aldine, 1967).

[76]Joseph L. Vigilante, "Between Values and Science: Education for the Profession During a Moral Crisis or Is Proof Truth", Journal of Education for Social Work, 10 (3), Fall 1974, pp. 107-115.

[77]Vigilante, "Between Values", p. 112.

[78]For further discussion, see: James S. Ackerman, "Two Styles: A Challenge to Higher Education", Daedalus, 98 (3), Summer 1969, pp. 855-869.

[79]Vigilante, "Between Values", p. 114.

[80]Vigilante, "Between Values", p. 113.

[81]E. Goffman, Asylyms: Essays in the Social Situation of Mental Patients and Other Inmates (New York: Anchor Books, 1961); and W. F. Whyte, Street Corner Society, second edition (Chicago: University of Chicago Press, 1969).

[82]H. S. Becker, B. Geer, and E. C. Hughes, Boys in White (Chicago: University of Chicago Press, 1961); C. Miller and M. Parlett, "Up to the Mark--a Research Report on Assessment", Occasional Paper 13, Centre for Research in the Educational Sciences, University of Edinburgh, January 1973; may also be found as: C. Miller and M. Parlett, Up to the Mark; a Study of the Examination Game, Monograph 21 (London: Society for Research into Higher Education, 1974).

[83]Clifford E. Brennan and Morton L. Arkava, "Students View the Undergraduate Curriculum", Journal of Education for Social Work, 10 (3), Fall 1974, pp. 9-16.

[84]For an interesting empirically oriented justification of this assumption, see: John A. Crane, Interpersonal Competence and Performance in Social Work (Vancouver: University of British Columbia, 1974). From study of 292 graduate students in social work, Crane's major findings suggest that student characteristics upon entering the program of instruction are primary determinants of their performance as students. See also: Kenneth A. Hack, "A Study of Predictors of Success on a Social Work Course", British Journal of Social Work, 3 (2), 1973, pp. 189-196.

[85]Brennan and Arkava, pp. 12-13.

[86]Brennan and Arkava, p. 14.

[87]Brennan and Arkava, p. 15.

[88]Gillian Michael, "Content and Method in Fieldwork Teaching", unpublished doctoral dissertation, University of Edinburgh, Department of Social Administration, 1976.

[89]Michael, p. 114.

[90]E. A. Brawley, "A Systems Approach to the Evaluation of a Community College Program in Mental Health Work" in Approaches to Innovation in Social Work Education (New York: Council on Social Work Education, 1974), pp. 15-25.

[91]Brawley, pp. 18-19.

CHAPTER III

[1]N. L. Gage, "Paradigms for Research on Teaching" in N. L. Gage (ed.), Handbook of Research on Teaching, American Educational Research Association (Chicago: Rand McNally, 1963), p. 95.

[2]Gage, Handbook, pp. 95-96.

[3]These labels were derived from a very useful discussion by Parlett and Hamilton which outlines the second approach: see Malcolm Parlett and David Hamilton, "Evaluation as Illumination, A New Approach to the Study of Innovatory Programs", Occasional Paper, Centre for Research in the Educational Sciences, University of Edinburgh, 1972. Another useful description and analysis of the two paradigms is provided in: Michael Quinn Patton, Alternative Evaluation Research Paradigm, North Dakota Study Group on Evaluation (Grand Forks: University of North Dakota, 1975). For relevant discussion within the social work literature, see: Peter W. Chommie and Joe Hudson, "Evaluation of Outcome and Process", Social Work, November 1974, pp. 682-687.

[4]See, for example, R. W. Tyler, Constructing Achievement Tests (Columbus: Ohio University Press, 1934); E. R. Smith and R. W. Tyler, Appraising Student Progress (New York: Harper and Row, 1942); R. W. Tyler, Basic Principles of Curriculum Instruction: A Syllabus for Education 360 (Chicago: University of Chicago Press, 1955); R. W. Tyler (ed.), Educational Evaluation: New Roles, New Means (Chicago: University of Chicago Press, 1969).

[5]D. Rowntree, Educational Technology in Curriculum Development (London: Harper & Row, 1974), p. 5.

[6]W. C. Jones and E. J. Borgatta, "Methodology of Evaluation" in E. J. Mullen, J. R. Dumpson and Associates, Evaluation of Social Intervention (London: Jossey-Bass, 1972), p. 40.

[7]May Brodbeck, "Logic and Scientific Method in Research on Teaching" in Gage, Handbook, p. 51.

[8]See for example, S. Wiseman and D. Pidgeon, Curriculum Evaluation (Slough: National Foundation for Educational Research, 1972), Ch. 2; B. S. Bloom, J. T. Hastings and G. F. Madaus, Handbook on Formative and Summative Evaluation of Student Learning (New York: McGraw-Hill, 1971), Ch. 2; B. S. Bloom, "Testing Cognitive Ability and Achievement" in N. L. Gage, Handbook, Ch. 8.

[9]Rowntree, p. 28.

[10]B. S. Bloom (ed.), Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook I, Cognitive Domain (London: Longmans, Green & Co., 1956); and D. R. Krathwohl, B. S. Bloom and B. B. Masia, Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook II, Affective Domain (U.S.A.: Longman, 1964).

[11]Krathwohl, pp. 6-7.

[12]Rowntree, p. 33.

[13]Rowntree, pp. 34-36.

[14]R. F. Mager and C. Clark, "Explorations in Student-Controlled Instruction", Psychological Reports, 13, 1963, pp. 71-76.

[15]Rowntree, p. 34.

[16]For discussion of "comparative designs", see L. Geismar, "Thirteen EValuative Studies" in Mullen and Dumpson, p. 28.

[17]Bloom, "Testing", p. 379.

[18]For example, see: R. L. Ebel, Measuring Educational Achievement (Englewood Cliffs, N.J.: Prentice-Hall, 1965); R. L. Thorndike and E. Hagen, Measurement and Evaluation in Psychology and Education, 3rd edition (New York: John Wylie & Sons, 1969).

[19]This distinction is clearly presented in C. I. Chase, Measurement for Educational Evaluation (London: Addison-Wesley, 1974), pp. 93-100; and discussed in W. J. Popham (ed.), Criterion-Referenced Measurement (Englewood Cliffs, N.J.: Educational Technology Publications, 1971).

[20]Rowntree, pp. 69-70.

[21]See P. O. Johnson and R. W. B. Jackson, Modern Statistical Methods: Descriptive and Inductive (Chicago: Rand McNally, 1959), Ch. 1.

[22]W. J. Popham, Educational Statistics: Use and Interpretation (New York: Harper & Row, 1967), pp. 5-6.

[23]A useful general discussion was provided by Egon G. Guba and Daniel L. Stufflebeam, "Evaluation: The Process of Stimulating, Aiding, and Abetting Insightful Action", address delivered at Second National Symposium for Professors of Educational Research, Nov. 21, 1968, Columbus, Ohio, Evaluation Centre, College of Education, Ohio State University.

[24]See D. A. MacKay and T. O. Maguire, Evaluation of Instructional Programs (Alberta: Human Resources Research Council, 1971), p. 6.

[25]Malcolm Parlett, "Evaluating Innovations in Teaching" in H. J. Butcher and E. Rudd (eds.), Contemporary Problems in Higher Education (London: McGraw-Hill, 1972), p. 147.

[26]Rowntree, p. 37.

[27]Rowntree, p. 57.

[28]McKay and Maguire, p. 6.

[29]Rowntree, p. 37.

[30]M. Trow, "Methodological Problems in the Evaluation of Innovation" in M. C. Wittrock and D. E. Wiley (eds.), The Evaluation of Instruction: Issues and Problems (New York: Holt, Rinehart and Winston, 1970), p. 298.

[31]Rowntree, p. 30.

[32]Rowntree, p. 40.

[33]A highly stimulating discussion in this area may be found in Michael Scriven, "Prose and Cons About Goal-Free Evaluation" in W. J. Popham (ed.), Evaluation in Education: Current Approaches (Berkely: McCutchan 1974), pp. 34-67.

[34]Rowntree, p. 40.

[35]Parlett and Hamilton, p. 5.

[36]J. P. Gilbert, R. L. Light and F. Mosteller, "Assessing Social Innovations: An Empirical Base for Policy" in C. A. Bennett and A. A. Lumsdaine (eds.), Evaluation and and Experiment: Some Critical Issues in Assessing Social Programs (New York: Academic Press, 1975), p. 44.

[37]Gilbert, Light and Mosteller, p. 159.

[38]Parlett in Butcher and Rudd, p. 146.

[39]Geismar in Mullen and Dumpson, p. 28.

[40]Parlett and Hamilton, p. 6.

[41]Parlett and Hamilton, p. 7.

[42]Rowntree, pp. 39-40.

[43]Jones and Bergatta in Mullen and Dumpson, p. 49.

[44]Jones and Bergatta in Mullen and Dumpson, pp. 40-41.

[45]E. Litwak, "Comments on Professor Trow's Paper: Toward a Multimodel Theory of Evaluation" in Wittrock and Wiley, p. 305.

[46]Parlett and Hamilton, p. 7.

[47]Parlett and Hamilton, p. 7.

[48]Parlett in Butcher and Rudd, p. 150.

[49]Gilbert, Light and Mosteller, pp. 144-145.

[50]Trow in Wittrock and Wiley, p. 297.

[51]Carol H. Weiss, Evaluation Research: Methods for Assessing Program Effectiveness (Englewood Cliffs, N.J.: Prentice-Hall, 1972), p. 25.

[52]Rowntree, p. 38.

[53]For useful general discussions, see R. Naroll and R. Cohen (eds.), A Handbook of Method in Cultural Anthropology (New York: Columbia University Press, 1973), Parts I and III.

[54]Parlett in Butcher and Rudd, p. 155.

[55]This label is provided by Trow in Wittrock and Wiley. Further definition may be found in Parlett and Hamilton's paper.

[56]Parlett in Butcher and Rudd, p. 144.

[57]Parlett and Hamilton, pp. 10-11.

[58]Parlett and Hamilton, p. 11.

[59]Parlett and Hamilton, p. 14.

[60]D. Jenkins, "Integrated Studies Project" in Evaluation in Curriculum Development: Twelve Case Studies, Schools Council Research Studies (London: Macmillan, 1973), pp. 70-79.

[61]Parlett and Hamilton, p. 18.

[62]For a discussion of related procedures, see E. J. Webb et al., Unobstrusive Measures: Non-Reactive Research in Social Sciences (Chicago: Rand McNally, 1966).

[63]Weiss, p. 36.

[64]P. J. Pelto, Anthropological Research: The Structure of Inquiry (New York: Harper & Row, 1970), p. 30.

[65]For example, see P. W. Bridgeman, The Logic of Modern Physics (London: Macmillan, 1927), p. 5.

[66]J. P. Nettl, Political Mobilization: A Sociological Analysis of Methods and Concepts (London: Faber and Faber, 1967).

[67]J. Ford, Paradigms and Fairy Tales: An Introduction to the Science of Meanings, Vol. I (London: Routledge & Kegan Paul, 1975), p. 150.

[68]Parlett and Hamilton, p. 25.

[69]Parlett and Hamilton, pp. 24-25.

[70]Sir Francis Bacon, The New Oraganon and Related Writings, (First Published in 1620), (Indianapolis: Bobbs-Merrill, 1960), p. 22.

[71]J. Ford, Paradigms and Fairy Tales: An Introduction to the Science of Meanings, Vol. II (London: Routledge & Kegan Paul, 1975), p. 259.

[72]Parlett and Hamilton, p. 26.

[73]Parlett and Hamilton, p. 28.

[74]For a very useful discussion of sampling issues, see Ford, Paradigms, Vol. II, Ch. 15.

[75]Parlett in Butcher and Rudd, p. 151.

[76]See Pelto, p. 42.

[77]R. Cohen, "Generalizations in Ethnology" in Naroll and Cohen, p. 34.

[78]For a discussion of "deep structures", see: Noam Chomsky, Syntactic Structures (The Hague: Mouton, 1957); and A. V. Cicourel, The Social Organization of Juvenile Justice (New York: Wiley, 1968), pp. 332-333.

[1]Allen Pincus and Anne Minahan, Social Work Practice: Model and Method (Itasca, Ill.: Peacock, 1973); Howard Goldstein, Social Work Practice: A Unitary Approach (Columbia: University of South Carolina Press, 1973); Ruth Middleman and Gale Goldberg, Social Service Delivery: A Structural Approach to Social Work Practice (New York: Columbia University Press, 1974); James K. Whittaker, Social Treatment: An Approach to Interpersonal Helping (Chicago: Aldine, 1974).

[2]Tutors' names were included in the report to the NISW.

[3]The "school-based" group included all social work teachers and teacher/administrators. The "agency-based" group included all others, except the fieldwork instructor who appeared to fit into neither category.

[4]The "amount learned" index was a combination of Items 2, 8a, 8b, 8c, and 8d on the "Overall Evaluation" instrument.

[1]Tutors' names were included in the report submitted to the NISW.

[2]For a relatively thorough conceptualization of this approach to evaluation, see Malcolm Parlett and David Hamilton, "Evaluation as Illumination: A New Approach to the Study of Innovatory Programs", Occasional Paper, Centre for Research in the Educational Sciences, University of Edinburgh, Edinburgh, October, 1972.

[3]"Direct practitioners" were defined as those who spent at least fifty percent of their time on the job working directly with or for "clients".

[4]The suggested textbooks were: Allen Pincus and Anne Minahan, Social Work Practice: Model and Method (Itasca, Ill.: F.E. Peacock, 1973); Howard Goldstein, Social Work Practice: A Unitary Approach (Columbia: University of South Carolina Press, 1973); Ruth Middleman and Gale Goldberg, Social Service Delivery: A Structural Approach to Social Work Practice (New York: Columbia Press, 1974); and James K. Whittaker, Social Treatment: An Approach to Interpersonal Helping (Chicago: Aldine, 1974).

CHAPTER VI

[1]Benjamin S. Bloom (ed.), Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 1, Cognitive Domain (London: Longmans, Green & Co., 1956).

[2]P. H. Rossi and W. Williams, Evaluating Social Programs (New York: Seminar Press, 1972), p. 18.

[3]Michael Scriven, "Prose and Cons About Goal-Free Evaluation" in W. James Popham (ed.), Evaluation in Education: Current Approaches (Berkeley: McCutchan, 1974), pp. 34-67.

[4]Carl A. Bennett and Arthur A. Lumsdaine, "Social Program Evaluation: Definitions and Issues" in C. A. Bennett and A. A. Lumsdaine (eds.), Evaluation and Experiment: Some Critical Issues in Assessing Programs (New York: Academic Press, 1975), p. 28.

[5]See, for example, the attempt to measure "motivation" in James Wilson, "Evaluation of Learning in Secondary School Mathematics" in B. Bloom, J. T. Hastings and G. F. Madeus, Handbook of Formative and Summative Evaluation of Student Learning (New York: McGraw-Hill, 1971), pp. 687-688.

[6]"Praxis" implies behavior which is purposive and accountable for only in terms of its meaningfulness. For interesting discussion, see: R. D. Laing and D. Cooper, Reason and Violence: A Decade of Sartre's Philosophy 1950-1960 (London: Tavistock, 1964); and David Ingleby, "Ideology of the Human Sciences: Some Comments on the Role of Reification in Psychology and Psychiatry" in Trevor Pateman (ed.), Counter Course: A Handbook of Course Criticism (Harmondsworth, Middlesex: Penguin, 1972), pp. 51-81. A useful perspective on issues related to motivation may also be found in R. M. Gagne, The Conditions of Learning, 2nd edition (New York: Holt, Rinehart & Winston, 1970).

[7]Hilda Taba, Curriculum Development: Theory and Practice (New York: Harcourt, Brace & World, 1962), p. 82.

[8]Carol H. Weiss, Evaluation Research: Methods for Assessing Program Effectiveness (Englewood Cliffs, N.J.: Prentice-Hall, 1972), p. 70.

[9]It must not be automatically assumed that selection procedures in social work education are effective. In that regard, interesting observations have been provided by:

John Kidneigh and Horace W. Lundberg, "Are Social Work Students Different?", Social Work, 3, (2), 1958, pp. 57-61; S. Berengarten, Admissions Prediction and Student Performance in Social Work Education (New York: Council on Social Work Education, 1964); B. F. Esser and C. Freud, Testing as a Tool in the Admissions Process: An Overview of Research (New York: Council on Social Work Education, 1966); and T. M. Brigham, "Philosophical Issues and Trends in Student Selection Today", Journal of Education for Social Work, 4, (2), Fall 1968, pp. 27-34.

[10]An issue of the journal Daedalus has been devoted to examination of ethical considerations in research with human subjects. See Daedalus, 98, (2), Spring 1969, Especially worthy of note are: Paul A. Freund, "Ethical Aspects of Experimentation on Human Subjects", pp. vii - xiv; Hans Jonas, "Philosophical Reflections on Experimenting With Human Subjects", pp. 219-247; and Talcott Parsons, "Research With Human Subjects and the Professional Complex", pp. 325-360.

[11]One should not discount the idea that well-intentioned instructional "interventions" might have harmful effects. If it is valid to assume that an instructional sequence may result in outcomes interpreted as "beneficial", it is equally valid to assume that instruction may have detrimental effects. For related commentary, see: Peter C. Loeb and John A. Slosar Jr., "Sociatrogenic Dysfunctions: A Concern for Social Work Education", Journal of Education for Social Work, 10 (2), Spring 1974, pp. 51-58; and Jerald G. Bachman, "Comment: Anti-Dropout Campaign and other Misanthropies", Society, 9 (5), March 1972.

[12]R. V. G. Clark and D. B. Cornish, The Controlled Trial in Institutional Research--Paradigm or Pitfall for Penal Evaluators, Home Office Research Unit (London: HMSO, 1972), p. 5.

[13]Clark and Cornish, p. 6.

[14]Ward Edwards and Marcia Guttentag, "Experiments and Evaluations: A Re-examination" in Bennett and Lumsdaine, Evaluation and Experiment, p. 457.

[15]James S. Ackerman, "Two Styles: A Challenge to Higher Education", Daedalus, 98 (3), Summer 1969, p. 856.

[16]Barney G. Glaser and Anselm L. Strauss, The Discovery of Grounded Theory: Strategies for Qualitative Research (Chicago: Aldine, 1967), p. 201.

[17]Scriven warns of various means by which educationally insignificant findings may acquire statistical significance. See: Michael Scriven, "Evaluation Perspectives and Procedures" in Popham, Evaluation in Education, p. 17.

[18]Glaser and Strauss, p. 243.

[19]James L. Breedlove, "Theory Development as a Task for the Evaluator" in E. J. Mullen, J. R. Dumpson and Associates, Evaluation of Social Intervention (London: Jossey-Bass, 1972), p. 58.

[20]Michael Q. Patton, In Search of Impact: An Analysis of the Utilization of Federal Health Evaluation Research, Minnesota Center for Social Research (Minneapolis: University of Minnesota, 1975), pp. 21-22.

[21]Egon G. Guba, "Problems in Utilizing the Results of Evaluation", Journal of Research and Development in Education, 8 (3), 1975, p. 45.

[22]Patton, p. 38.

[23]Kenneth O. Doyle Jr., Student Evaluations of Instruction (Lexington, Mass.: D. C. Heath & Co., 1975), p. 65.

[24]Doyle, Student Evaluations, p. 22.

[25]See, for example: W. B. Webb and C. Y. Nolan, "Student, Supervisor and Self-Rating of Instructional Proficiency", Journal of Educational Psychology, 46, 1955, pp. 42-46; R. T. Blackburn and M. Clark, "An Assessment of Faculty Performance: Some Correlates Between Administrator, Colleague, Student, and Self-Ratings", Ann Arbor Center for the Study of Higher Education, University of Michigan, 1971; J. A. Centra, "Two Studies on the Utility of Student Ratings for Improving Teaching", SIR Report No. 2, Princeton, N.J., Educational Testing Service, 1972.

[26]See, for example: A. J. Drucker and H. H. Remmers, "Do Alumni and Students Differ in Their Attitudes Toward Instructors?", Journal of Educational Psycholoy, 42, 1951, pp. 129-143; A. H. Maslow and W. Zimmerman, "College Teaching Ability, Scholarly Activity and Personality", Journal of Educational Psychology, 47, 1956, pp. 185-189; J. A. Centra, "The Relationship Between Student and Alumni Ratings of Teachers", Research Bulletin 73-79, Princeton, N.J., Educational Testing Service, 1973; L. M. Aleamoni and M. Yimer, "An Investigation of the Relationship Between

Colleague Rating, Student Rating, Research Productivity, and Academic Rank in Rating Instructional Effectiveness", Journal of Educational Psycholoy, 64, 1973, pp. 274-277; J. Centra, "College Teaching: Who Should Evaluate It?" Findings, 1, (1), Princeton, N.J.: Educational Testing Service, 1974.

[27]See, for example: A. W. Bendig, "Student Achievement in Introductory Psychology and Student Ratings of the Competence and Empathy of Their Instructors", Journal of Educational Psychology, 36, 1953, pp. 427-433; G. D. Lovell and C. F. Haner, "Forced-Choice Scales Applied to College Faculty Rating", Educational and Psychological Measurement, 15, 1955, pp. 291-304; N. F. Rayder, "College Student Ratings of Instructors", Journal of Experimental Education, 37, (2), 1968, pp. 76-81; K. O. Doyle Jr., I. I. Bejar, and S. E. Whitely, "Norms for the Student Opinion Survey", Measurement Services Center, Minneapolis, University of Minnesota, 1974; K. O. Doyle Jr. and S. E. Whitely, "Student Ratings as Critera for Effective Teaching", American Educational Research Journal, 11, (3), 1974, pp. 259-274.

[28]A. T. Sharon, "Eliminating Bias from Student Ratings of College Instructors", Journal of Applied Psychology, 54, 1970, pp. 278-281; A. T. Sharon and C. J. Bartlett, "Effect of Instructional Conditions in Producing Leniency on Two Types of Rating Scales", Personnel Psychology, 22, 1969, pp. 251-263.

[29]Doyle, Student Evaluations, pp. 81-82.

[30]W. J. McKeachie, "Student Ratings of Faculty", AAUP Bulletin, 55, 1969, pp. 439-444; Centra, "Two Studies"; L. M. Alemoni, "The Usefulness of Student Evaluations in Improving College Teaching", Urbana, Measurement and Research Division, Office of Instructional Resources, University of Illinois, 1974.

[31]Neely undertook a twenty-five year review of the literature on criteria for the identification of effective teaching and found no objective usable criteria; see M. Neely, "A Teacher's View of Teacher Evaluation", Improving College and University Teaching, 16, 1968, pp. 207-209.

CHAPTER VII

[1]Malcolm Parlett, "Evaluating Innovations in Teaching" in H. J. Butcher and E. Rudd (eds.), Contemporary Problems in Higher Education (London: McGraw-Hill, 1972), p. 151.

[2]Peter Leonard, "Explanation and Education in Social Work", British Journal of Social Work, 5, (3), 1975, p. 332.

[3]Malcolm Parlett and David Hamilton, "Evaluation as Illumination, A New Approach to the Study of Innovatory Programs", Occasional Paper, Centre for Research in the Educational Sciences, University of Edinburgh, 1972, p. 39.

[4]For further clarification of this difference in sample types, see the discussion of "Theoretical Sampling" in Barney G. Glaser and Anselm L. Strauss, The Discovery of Grounded Theory: Strategies for Qualitative Research (Chicago: Aldine, 1967), pp. 61-63.

[5]Donald T. Campbell and Robert F. Borush, "Making the Case for Randomized Assignment to Treatments by Considering the Alternatives: Six Ways in Which Quasi-experimental Evaluations in Compensatory Education Tend to Underestimate Effects" in C. A. Bennett and A. A. Lumsdaine (eds.) Evaluation and Experiment: Some Critical Issues in Assessing Programs (New York: Academic Press, 1975), pp. 199-200.

[6]James S. Ackerman, "Two Styles: A Challenge to Higher Education", Daedaulus, 98, (3), Summer 1969, p. 860.

[1]Carl A. Bennett and Arthur A. Lumsdaine, "Social Programe Evaluation: Definitions and Issues" in C. A. Bennett and A. A. Lumsdaine (eds.), Evaluation and Experiment: Some Critical Issues in Assessing Programs (New York: Academic Press, 1975), p. 1.

[2]W. Reid and A. Shyne, Brief and Extended Casework (New York: Columbia University Press, 1969), p. 55.

[3]For excellent discussion, see: Richard M. Wolf, "Data Analysis and Reporting Considerations in Evaluation" in W. James Popham (ed.), Evaluation in Education: Current Applications (Berkeley: McCutchan, 1974), pp. 203-242.

[4]A useful set of recommendations for assessments of program impact is contained in: Elizabeth T. Hilton, "Field Trial Designs in Gauging the Impact of Fertility Planning Programs" in Bennett and Lumsdaine (eds.), Evaluation and Experiment, pp. 319-408.

[5]Margaret Mead, "Research with Human Beings: A Model Derived from Anthropological Field Practice", Daedalus, 98, (2), Spring 1969, p. 361.

[6]Gillian Michael, Content and Method in Fieldwork Teaching, unpublished doctoral dissertation, University of Edinburgh, 1976, p. 533.

[7]Gillian Michael, p. 201.

[8]Martin Davies and Elinor Kelly, "The Social Worker, the Client and the Social Anthropologist", British Journal of Social Work, 6, (2), 1976, p. 214-231.

[9]Michael's observations with regard to her interview pro forma closely correspond with my conclusions about the questionnaire and interview schedule used in Study II. See Gillian Michael, p. 445.

[10]Michael, for example, observed that these consultations were valuable in her study. See Gillian Michael, p. 208.

[11]Steven Wiseman and Douglas Pidgeon, Curriculum Evaluation (Slough, Bucks.: National Foundation for Educational Research, 1972), p. 29.

[12]M. Parlett and D. Hamilton, "Evaluation as Illumination: A New Approach to the Study of Innovatory Programs", Occasional Paper, Centre for Research in the Educational Sciences, University of Edinburgh, 1972, p. 9.

[13]See discussions by Martin Trow and Eugene Litwak in M. C. Wittrock and D. E. Wiley (eds.), The Evaluation of Instruction: Issues and Problems (New York: Holt, Rinehart and Winston, 1970), Ch. 9, pp. 305-317.

[14]John Carrier and Ian Kendall, "Social Policy and Social Change--Explanations of the Development of Social Policy", Journal of Social Policy, 2, (3), 1973, p. 212.

[15]For an excellent discussion of this position as it applies to psychological and psychiatric endeavours, see: David Ingleby, "Ideology and the Human Sciences: Some Comments on the Role of Reification in Psychology and Psychiatry" in Trevor Pateman (ed.), Counter Course: A Handbook for Course Criticism (Harmondsworth: Penguin, 1972), pp. 51-81.

[16]See, for example: T. Kuhn, The Structure of Scientific Revolution (Chicago: Chicago University Press, 1962); R. D. Laing and D. Cooper, Reason and Violence: A Decade of Sartre's Philosophy 1950-1960 (London: Tavistock, 1964); Peter L. Berger and Thomas Luckman, The Social Construction of Reality (Harmondsworth: Penguin, 1967); and L. Goldman, The Human Sciences and Philosophy, trans. by H. V. White and R. Anchor, Jonathan Cape, 1969.

[17]C. Guzetta, "Curriculum Alternatives", Journal of Education for Social Work, 8, (2), Winter 1972, p. 25.

[18]B. Rothman and J. L. Vigilante, "Curriculum Planning in Social Work Education", Journal of Education for Social Work, 10 (2), Spring 1974, pp. 76-85.

[19]Rothman and Vigilante, p. 80.

[20]Andrew's approach to predicting social work practice needs and Vinter and Seigel's formulation of operational areas of data need may represent useful approaches in this area. See: Gwen Andrew, "Forecasting Social Work Practice as a Base for Curriculum Development", Journal of Education for Social Work, 10 (3), Fall 1974, pp. 3-8; and R. D. Vinter and S. Seigel, "Proposal for Schools of Social Work to Establish a National Information System", Ann Arbor, Michigan, 1971.

[21]For relevant discussion, see: Ann L. McLean and Joan F. Robertson, "Systematic Integration of Consumer Input into Curriculum Revision", Journal of Education for Social Work, 11 (2), Spring 1975, pp. 84-90.

BIBLIOGRAPHY

BIBLIOGRAPHY

VOLUMES

American Institutes for Research; Evaluative Research: Strategies and Methods (Pittsburgh: 1970).

Backstrom C.H., and G. D. Hursh; Survey Research (Evanston, Ill.: Northwestern University Press, 1963).

Bacon, Sir Francis; The New Organon and Related Writings, (First Published in 1620), (Indiannapolis: Bobbs-Merrill, 1960).

Barnette, W. L. Jr. (ed.); Readings in Psychological Tests and Measurements (Homewood, Ill.: Dorsey, 1968).

Becker, H. S., B. Geer, and E. C. Hughes; Boys in White (Chicago: University of Chicago Press, 1961).

Belasco, J. A. and H. M. Trice; The Assessment of Change in Training and Therapy (New York: McGraw-Hill, 1969).

Bennett, Carl A. and Arthur A. Lumsdaine (eds.); Evaluation and Experiment: Some Critical Issues in Assessing Social Programs (New York: Academic Press, 1975).

Bennis, W. G., K. D. Benne and R. Chin (eds.); The Planning of Change (New York: Holt, Rinehart, and Winston, 1964).

Berdie, D. R. and J. F. Anderson; Questionnaires: Design and Use (Metuchen, N.H.: Scarecrow Press, 1974).

Berengarten, S.; Admissions Predictions and Student Performance in Social Work Education (New York: Council on Social Work Education, 1964).

Berger, Peter L. and Thomas Luckman; The Social Construction of Reality (Harmondsworth: Penguin, 1967).

Biddle, Bruce J. and William J. Ellena; Contemporary Research on Teacher Effectiveness (New York: Holt, Rinehart, and Winston, 1964).

Bilodeau, E. A.; Acquisition of Skill (New York: Academic Press, 1966).

Blackey, Eileen A.; Group Leadership in Staff Training (Washington D.C.: Department of Health, Education, and Welfare, Bureau of Family Services, 1956).

Blalock, Hubert M. Jr., Causal Inferences in Nonexperimental Data (Chapel Hill: University of North Carolina Press, 1964).

Bloom, Benjamin S. (ed.); Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 1, Cognitive Domain (London: Longmans, Green & Co., 1956).

Bloom, Benjamin S., J. Thomas Hastings, and George F. Madaus; Handbook on Formative and Summative Evaluation of Student Learning (New York: McGraw-Hill, 1971).

Bloom, Martin and Marcella Farrar; Final Report: Becoming a Social Worker (Cleveland: Benjamin Rose Institute, 1970).

Bridgeman, P. W.; The Logic of Modern Physics (London: Macmillan, 1927).

Bruner, J. S.; The Process of Education (Cambridge, Mass.: Harvard University Press, 1960).

_____. On Knowing: Essays for the Left Hand (New York: Atheneum, 1967).

_____., R. R. Oliver, P. M. Greenfield, et al.; Studies in Cognitive Growth (New York: John Wylie & Sons, 1966).

Bruyn, S.; The Human Perspective in Sociology: The Methodology of Participant Observation (Englewood Cliffs, N.J.: Prentice-Hall, 1966).

Butcher, H. J. and E. Rudd (eds.); Contemporary Problems in Higher Education (London: McGraw-Hill, 1972).

Campbell, Donald T. and Julian C. Stanley; Experimental and Quasi-experimental Designs for Research (Chicago: Rand-McNally, 1963).

Caro, Francis G. (ed.); Readings in Evaluative Research (New York: Russell Sage Foundation, 1971).

Carter, Novia and Brian Wharf; Evaluating Social Development Programs (Ottawa: Canadian Council on Social Development, 1973).

Chase, Clinton I.; Measurement for Educational Evaluation (Reading, Mass.: Addison-Wesley, 1974).

Cherney, Paul R. (ed.); Making Evaluation Research Useful (Columbia, Maryland: American City Corporation, 1971).

Chomsky, Noam; Syntactic Structures (The Hague: Mouton, 1957).

Chronbach, Lee J.; Essentials of Psychological Testing, third edition (New York: Harper & Row, 1970).

Churchman, C. W.; The Systems Approach (New York: Delacorte Press, 1968).

Cicourel, A. V.; The Social Organization of Juvenile Justice (New York: Wiley, 1968).

Clark, R. V. G. and D. B. Cornish; The Controlled Trial in Institutional Research--Paradigm or Pitfall for Penal Evaluators, Home Office Research Unit (London: HMSO, 1972).

Compeau, P. L.; Selective Review of the Results of Research on the Use of Audiovisual Media to Teach Adults (Strasbourg: Council of Europe, 1972).

Council on Social Work Education; Curriculum Study, Vols. 1-12 (New York: CSWE, 1959).

Council on Social Work Education; A Source Book of Readings on Teaching in Social Work: Reprints of Selected Articles (New York: CSWE, 1965).

Council on Social Work Education; Approaches to Innovation in Social Work Education (New York: CSWE, 1974).

Crane, John A.; Interpersonal Competence and Performance in Social Work (Vancouver: University of British Columbia, 1974).

Davis, Frederick B.; Educational Measurements and Their Interpretation (Belmont, Calif.: Wadsworth, 1964).

Dea, K. (ed.); New Ways of Teaching Social Work Practice (New York: Council on Social Work Education, 1972).

Doyle, Kenneth O. Jr.; Student Evaluation of Instruction (Lexington, Mass.: D. C. Heath & Co., 1975).

Dubin, R. and T. C. Taveggia; The Teaching-Learning Paradox: A Comparative Analysis of College Teaching Methods (Portland, Oregon: University of Oregon, 1968).

Dubois, Philip H. and E. Douglas Mayer (eds.); Research Strategies for Evaluating Training, AERA Monograph Series on Evaluation, No. 4 (Chicago: Rand McNally, 1970).

Ebel, R. L.; Measuring Educational Achievement (Englewood Cliffs, N.J.: Prentice-Hall, 1965).

Ebel, R. L (ed.); Encyclopedia of Educational Research, 4th edit. (New York: Macmillan, 1969).

Esser, B. F. and C. Freud; Testing as a Tool in the Admissions Process: An Overview of Research (New York: Council on Social Work Education, 1966).

Fiegel, H. and May Broadbeck (eds.); Readings in the Philosophy of Science (Appleton, Century-Crofts, 1953).

Fischer, Joel; The Effectiveness of Social Casework (Springfield, Ill.: Charles C. Thomas, 1976).

Ford, J.; Paradigms and Fairy Tales: An Introduction to the Science of Meanings, Vols. I & II, (London: Routledge & Kegan Paul, 1975).

Gage, N. L. (ed.); Handbook of Research on Teaching (Chicago: Rand-McNally, 1963).

Gagne, R. M.; The Conditions of Learning, 2nd edition (New York: Holt, Rinehart & Winston, 1970).

Glaser, Barney G. and Anselm L. Strauss; The Discovery of Grounded Theory: Strategies for Qualitative Research (Chicago: Aldine, 1967).

Glass, G. V.; The Growth of Evaluation Methodology, AERA Monograph Series on Curriculum Evaluation (Chicago: Rand McNally, 1971).

Glass, Gene V. (ed.); Evaluation Studies Review Annual, Vol. 1 (Beverley Hills: Sage Publications, 1976).

Glick, L. (ed.); Undergraduate Social Work Education for Practice: A Report on Curriculum Content and Issues, Syracuse University School of Social Work (Washington, D.C.: U.S. Government Printing Office, 1971).

Gnagy, W. J., P. A. Chesebro, and J. A. Johnson; Learning Environments (New York: Rinehart and Winston, 1972).

Goffman, E.; Asylums: Essays in the Social Situation of Mental Patients and Other Inmates (New York: Anchor Books, 1961).

Goldberg, E. Matilda, Ann Mortimer and Brian T. Williams; Helping the Aged: A Field Experiment in Social Work (London: George Allen & Unwin, 1970).

Goldman, L.; The Human Sciences and Philosophy, trans. by H. V. White and R. Anchor (London: Jonathan Cape, 1969).

Goldstein, Howard; Social Work Practice: A Unitary Approach (Columbia: University of South Carolina Press, 1973).

Gronlund, N. E.; Measurement and Evaluation in Teaching (New York: Macmillan, 1971).

Guilford, J. P.; Psychometric Methods (New York: McGraw-Hill, 1954).

Guilford, J. P.; The Nature of Human Intelligence (New York: McGraw-Hill, 1967).

Guttentag, M. and E. L. Streuning (eds.); Handbook of Evaluation Research, Vol. 2 (Beverley Hills: Sage, 1975).

Guttentag, Marcia and Shalom Saar (eds.); Evaluation Studies Review Annual, Vol. 2 (Beverley Hills: Sage Publications, 1977).

Helmstadter, G. C.; Principles of Psychological Measurement (New York: Appleton, Century-Crofts, 1964).

Herzog, Allen and L. R. Denton; A Compendium of Papers on Evaluation and Methodology (Yarmouth, Nova Scotia: Nova Scotia Newstart Inc., 1971).

Herzog, Elizabeth; Some Guidelines for Evaluative Research: Assessing Psycho-Social Change in Individuals (Washington D.C.: Department of Health, Education, and Welfare, Children's Bureau, 1959).

Hessling, P.; Strategy of Evaluation Research in the Field of Supervisory and Management Training (Assen, Netherlands: Van Gorcum, 1966).

Heyman, Margaret; Criteria and Guidelines for the Evaluation of In-Service Training (Washington D.C.: Department of Health, Education and Welfare, 1967).

Heywood, J. and A. H. Illife; Some Problems of Testing Academic Performance (Department of Higher Education, University of Lancaster, 1966).

Hoyt, Donald P.; The Relationship Between College Grades and Adult Achievement (Iowa City, Iowa: American College Testing Program Research Report No. 7, 1965).

Hudson, B.; Assessment Techniques (London: Methuen, 1973).

Hudson, L.; The Cult of the Fact (London: Jonathan Cape, 1972).

Jeffreys, H.; Scientific Inference (Cambridge University Press, second edition, 1937).

Johnson, P. O., and R. W. B. Jackson; Modern Statistical Methods: Descriptive and Inductive (Chicago: Rand McNally, 1959).

Klein, Alan F.; Role Playing in Leadership Training and Problem Solving (New York: Association Press, 1956).

Krathwohl, D. R., B. S. Bloom, and B. B. Masia; Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook II, Affective Domain USA: Longmans, 1964).

Kuhn, Thomas H.; The Structure of Scientific Revolutions (Chicago: University of Chicago Press, 1965).

Labovitz, Sanford and Robert Hagedorn; Introduction to Social Research (New York: McGraw-Hill, 1971).

Laing, R. D. and D. Cooper; Reason and Violence: A Decade of Sartre's Philosophy 1950-1960 (London: Tavistock, 1964).

Lindquist, E. F. (ed.); Educational Measurement (Washington: American Council on Education, 1951).

Loewenberg, Frank M.; Time and Quality in Graduate Social Work Education (New York: Council on Social Work Education, 1972).

Lopez, F. M.; Evaluating Executive Decision Making (New York: American Management Association, 1966).

Lynton, Rolf, P. and Udai Pareek; Training for Development (Illinois: Irwin-Dorsey, 1967).

Maas, H. S. (ed.); Five Fields of Social Service: Reviews of Research (New York: National Association of Social Workers, 1966).

_____. Research in the Social Services: A Five Year Review (New York: National Association of Social Workers, 1971).

MacKay, D. A. and T. O. Maguire; Evaluation of Instructional Programs (Edmonton, Alberta: Human Resources Research Council, 1971).

Mager, R. F.; Preparing Instructional Objectives (Palo Alto, Calif.: Fearon, 1962).

McKeachie, W. J.; Teaching Tips: A Guidebook for the Beginning College Teacher, 6th edition. (Lexington, Mass.: D. C. Heath & Co., 1969).

Mehren, W. A. and I. J. Lehmann; Standaridized Tests in Education (New York: Holt, Rinehart & Winston, 1969).

Merril, M. D. (ed.); Instructional Design: Readings (Englewood Cliffs, N.J.: Prentice-Hall, 1971).

Michael, Gillian; "Content and Method in Fieldwork Teaching", unpublished doctoral dissertation, University of Edinburgh, Department of Social Administration, 1976.

Middleman, Ruth and Gale Goldberg; Social Service Delivery: A Structural Approach to Social Work Practice (New York: Columbia University Press, 1974).

Miller, C. and M. Parlett; Up to the Mark: A Study of the Examination Game, Monograph 21 (London: Society for Research into Higher Education, 1974).

Miller, Richard I.; Evaluating Faculty Performance (London: Jossey-Bass, 1972).

Moscrop, Martha; In Service Training for Social Agency Practice (Toronto: University of Toronto Press, 1958).

Mullen, E. J., J. R. Dumpson, and Associates; Evaluation of Social Intervention (London: Jossey-Bass, 1972).

Naroll, Raoul and Ronald Cohen (eds.); A Handbook of Method in Cultural Anthropology (New York: Columbia University Press, 1970).

National Association of Social Workers; Guidelines for the Assessment of Professional Competence in Social Work, Committee on the Study of Competence (New York: NASW, 1968).

Nettl, J. P.; Political Mobilization: A Sociological Analysis of Methods and Concepts (London: Faber and Faber, 1967).

Nunnally, J. C.; Psychometric Theory (New York: McGraw-Hill, 1967).

Oppenheim, A. N.; Questionnaire Design and Attitude Measurement (London: Heinemann 1966).

Osgood, C. E., G. J. Suci, and P. H. Tannenbaum; The Measurement of Meaning (Urbana, Ill.: University of Illinois, 1957).

Oviatt, Boyd E. (ed.); Evaluation and Accountability in Social Work Education (Salt Lake City: Graduate School of Social Work, University of Utah, 1973).

Pateman, Trevor (ed.); Counter Course: A Handbook for Course Criticism (Harmondsworth, Middlesex: Penguin, 1972).

Patton, Michael Q. In Search of Impact: An Analysis of the Utilization of Federal Health Evaluation Research, Minnesota Center for Social Research (Minneapolis: University of Minnesota, 1975), pp. 21-22.

Patton; Michael Quinn; Alternative Evaluation Research Paradigm, North Dakota Study Group on Evaluation (Grand Forks: University of North Dakota, 1975).

Payne, D. A.; The Specification and Measurement of Learning Outcomes (Waltham, Mass.: Blaisdell, 1965).

Pelto, Pertti J.; Anthropological Research: The Structure of Enquiry (New York: Harper & Row, 1970).

Pettes, Dorothy; Supervision in Social Work (London: George Allen & Unwin, 1967).

Pincus, Allen and Anne Minahan; Social Work Practice: Model and Method (Itasca, Ill.: Peacock, 1973).

Pins, Arnulf; Who Chooses Social Work, When and Why? (New York: Council on Social Work Education, 1963).

Platt, William J.; Research for Educational Planning: Notes on Emergent Needs, UNESCO, International Institute for Educational Planning, 1970.

Plowman, P. D.; Behavioral Objectives (Chicago: Science Research Associates, 1971).

Pohek, Marguerite (ed.); The Teacher's Compendium (New York: Council on Social Work Education, 1963).

Pohek, Marguerite (ed.); Teaching and Learning in Social Work Education (New York: Council on Social Work Education, 1970).

Polansky, Norman A.; Social Work Research (Chicago: University of Chicago Press, 1960).

Popham, W. J.; Educational Statistics: Use and Interpretation (New York: Harper & Row, 1967).

Popham, W. James; Instructional Objectives, AERA Monograph Series on Curriculum Evaluation, No. 3 (Chicago: Rand McNally, 1969).

Popham, W. J. (ed.); Criterion-Referenced Measurement (Englewood Cliffs, N.J.: Educational Technology Publications, 1973).

Popham, W. James (ed.); Evaluation in Education: Current Applications (Berkeley: McCutchan, 1974).

Pumphrey, Muriel; The Teaching of Values and Ethics in Social Work Education (Vol. XIII of the Curriculum Study (New York: Council on Social Work Education, 1959).

Reid, W. and A. Shyne; Brief and Extended Casework (New York: Columbia University Press, 1969).

Remmers, H. H. and N. L. Gage; Educational Measurement and Evaluation, Revised Edit. (New York: Harper, 1955).

Reynolds, Bertha Caplan; Learning and Teaching in the Practice of Social Work (New York: Farrar & Rinehart, 1942).

Ripple, Lillian (ed.); Innovations in Teaching Social Work Practice (New York: Council on Social Work Education, 1970).

Romiszowski, A. J.; The Systems Approach to Education and Training (Page: London, 1970).

Rosenshine, Barak; Teaching Behaviors and Student Achievement (Slough, Bucks.: National Foundation for Educational Research in England and Wales, 1971).

Rossi, Peter H. and Walter Williams (eds.); Evaluating Social Programs: Theory, Practice, and Politics (New York: Seminar Press, 1972).

Rowntree, D.; Educational Technology in Curriculum Development (London: Harper & Row, 1974).

Schools Council; Evaluation in Curriculum Development: Twelve Case Studies, Schools Council Research Studies (London: Macmillan, 1973).

Sellitz, Claire and Marie Jahoda; Research Methods in Social Relations, Revised Edit. (New York: Holt, 1959).

Shaw, M. E. and J. M. Wright; Scales for the Measurement of Attitudes (New York: McGraw-Hill, 1967).

Shulman, L. S., and E. R. Keislar; Learning by Discovery: A Critical Appraisal (Chicago: Rand-McNally, 1966).

Smith, E. R., and R. W. Tylor; Appraising and Recording Student Progress, The Progressive Education Association (New York: McGraw-Hill, 1942).

Soffen, Joseph; Faculty Development in Professional Education (New York: Council on Social Work Education, 1967).

Stanley, J. C., and K. D. Hopkins; Educational and Psychological Measurement and Evaluation (Englewood Cliffs, N.J.: Prentice-Hall, 1972).

Stones, E.; Educational Objectives and the Teaching of Educational Psychology (London: Methuen, 1972).

Stufflebeam, Daniel L. et al.; Educational Evaluation and Decision-Making (Itaska, Ill.: F. E. Peacock, 1971).

Streuning, E. L. and M. Guttentag; Handbook of Evaluation Research, Vol. 1 (Beverly Hills: Sage, 1975).

Suchman, Edward A.; Evaluative Research: Principles and Practice in Public Service and Social Action Programs (New York: Russell Sage Foundation, 1967).

Taba, Hilda; Curriculum Development: Theory and Practice (New York: Harcourt, Brace & World, 1962).

Taylor, L. E.; Tests and Measurements, second edition (Englewood Cliffs, N.J.: Prentice-Hall, 1971).

Thorndike, R. L. (ed.); Educational Measurement (Washington, D.C.: American Council on Education, 1971).

Thorndike, Robert L. and Elizabeth Hagen; Measurement and Evaluation in Psychology and Education, 2nd edit. (New York: John Wiley & Sons, 1960).

Towle, Charlotte; The Learner in Education for the Professions (Chicago: University of Chicago Press, 1954).

Travers, Robert M. W.; How to Make Achievement Tests (New York: Odyssey Press, 1950).

Travers, R. M. W.; Essentials of Learning: An Overview for Students of Education, 2nd edit. (New York: Macmillan, 1967).

Travers, R.M.W.; An Introduction to Educational Research, third edition (London: McMillan, 1969).

Travers, Robert M. W. (ed.); Second Handbook of Research on Teaching, American Educational Research Association (Chicago: Rand McNally, 1973).

Tripodi, Toni, Phillip Fellin, and Irwin Epstein; Social Program Evaluation: Guidelines for Health, Education and Welfare Administrators (Itasca, Ill.: F. E. Peacock, 1971).

Tyler, R. W.; Constructing Achievement Tests (Columbus: Ohio University Press, 1934).

Tyler, R. W.; Basic Principles of Curriculum Instruction: A Syllabus for Education 360 (Chicago: University of Chicago Press, 1955).

Tyler, Ralph; Building the Social Work Curriculum (New York: Council on Social Work Education, 1960).

Tyler, R. W., (ed.); Educational Evaluation: New Roles, New Means, 68th Year Book of the National Society for the Study of Education (Chicago: University of Chicago Press, 1969).

Tyler, R. W., R. Gagne, and M. Scriven; Perspectives of Curriculum Evaluation (Chicago: Rand McNally, 1967).

U. S. Department of Health, Education and Welfare; Preparing Evaluation Reports: A Guide for Authors (Washington, D.C.: Government Printing Office, 1970).

Verreck, W. A. (ed.); Methodological Problems in Research and Development in Higher Education (Amsterdam: Swets & Zeitlinger, 1974).

Wahlberg, H. J. (ed.); Evaluating Educational Performance (Berkeley: McCutchan, 1974).

Warren, Malcolm, W.; Training for Results (Reading, Mass.: Addison-Weseley, 1969).

Webb, E. J., et al; Unobstrusive Measures: Nonreactive Research in the Social Sciences (Chicago: Rand McNally, 1969).

Weiss, Carol H. (ed.); Evaluating Action Programs (Boston: Allyn & Bacon, 1972).

Weiss, Carol; Evaluation Research: Methods for Assessing Program Effectiveness (Englewood Cliffs, N.J.: Prentice-Hall, 1972).

Wheeler, D. K.; Curriculum Process (London: University of London Press, 1967).

Whitelaw, Matt; The Evaluation of Management Training: A Review (London: Institute of Personnel Management, 1972).

Whittaker, James K.; Social Treatment: An Approach to Interpersonal Helping (Chicago: Aldine, 1974).

Whyte, W. F.; Street Corner Society, second edition (Chicago: University of Chicago Press, 1969).

Wiseman, Steven and Douglas Pidgeon; Curriculum Evaluation (Slough, Bucks.: National Foundation for Educational Research, 1972).

Wittich, W. A., and C. F. Schuller; Instructional Technology: Its Nature and Use, 5th edit., (New York: Harper and Row, 1973).

Wittrock, M. C. and D. E. Wiley (eds.); The Evaluation of Instruction: Issues and Problems (New York: Holt, Rinehart and Winston, 1970).

Wood, D. A.; Test Construction (Columbus, Ohio: Charles E. Merril, 1960).

Younghusband, E.; Social Work and Social Values (London: Allen and Unwin, 1967).

ARTICLES AND PAPERS

Abt, M. H. and G. L. Fahey; "A Measurement of College Instructor Behavior", Scientia Paedagogica Experimentalis, 5, 1968, pp. 32-39.

Ackerman, James S.; "Two Styles: A Challenge to Higher Education", Daedalus, 98 (3), Summer 1969, pp. 855-869.

Aleamoni, L. M. and M. Yimer; "An Investigation of the Relationship Between Colleague Rating, Student Rating, Research Productivity, and Academic Rank in Rating Instructional Effectiveness", Journal of Educational Psychology, 64, 1974, 274-277.

Aleamoni, L. M.; "The Usefulness of Student Evaluations in Improving College Teaching", Urbana, Measurement and Research Division, Office of Instructional Resources, University of Illinois, 1974.

Andrew, Gwen; "Forecasting Social Work Practice as a Base for Curriculum Development", Journal of Education for Social Work, 10, (3), 1974, pp. 3-8.

Atkinson, P., M. Reid, and P. Sheldrake; "Medical Mystique: Interdeterminancy and Models of Professional Process", Occasional Paper No. 14, Centre for Research in the Educational Sciences, University of Edinburgh, February 1973.

Bachman, Jerald G. "Comment: Anti-Dropout Campaign and Other Misanthropies", Society, 9 (5), March 1972.

Ballard, R. G. and E. H. Mudd; "Some Sources of Difference Between Client and Agency Evaluation of Effectiveness of Counselling", Social Casework, 39 (1) 1958, pp. 30-35.

Barker, M. E.; "Education for Social Work, a Teacher's Viewpoint", Social Work Today, 1 (9), 1970), pp. 18-23.

Bateman, Worth; "Assessing Program Effectiveness: A Project Rating System Described", Welfare in Review, 6 (1), 1968, pp. 1-10.

Becker, H. S.; "Problems of Inference and Proof in Participant Observation", American Sociological Review, 23, 1958, pp. 652-660.

Bejar, I. I. and K. O. Doyle Jr.; "Expectations, First Impressions and Evaluations", Measurement Services Center, University of Minnesota, Minneapolis, 1974.

Bejar, I. I. and K. O. Doyle Jr.; "The Effect of Prior Expectation on the Structure and Evaluation of Student Ratings of Teacher Behavior", Measurement Services Center, University of Minnesota, Minneapolis, 1974.

Bell, Winifred; "Obstacles to Shifting from the Descriptive to the Analytic Approach in Teaching Social Services", Journal of Education for Social Work, 5 (1), 1969.

Bendig, A. E.; "Student Achievement in Introductory Psychology and Student Ratings of the Competence and Empathy of Their Instructors", Journal of Psychology, 36, 1953, pp. 427-433.

Bennett, Carl A. and Arthur A. Lumsdaine, "Social Program Evaluation: Definitions and Issues", in C. A. Bennett and A. A. Lumsdaine (eds.), Evaluation and Experiment: Some Critical Issues in Assessing Programs (New York: Academic Press, 1975), pp. 1-34.

Berger, M. and R. Williams; "Evaluation of Management Development, A Summary Report", Department of Occupational Psychology, Birkbeck College, London, February, 1973 (mimeo).

Berk, Richard A. and Peter Rossi; "Doing Good or Worse: Evaluation Research Politically Reexamined" in Marcia Guttentag and Shalom Saar (eds.), Evaluation Studies Review Annual (Beverley Hills: Sage, 1977), pp. 77-90.

Blackburn, R. T. and M. Clark; "An Assessment of Faculty Performance: Some Correlates Between Administrator, Colleague, Student, and Self-Ratings", Ann Arbor Center for the Study of Higher Education, University of Michigan, 1971.

Blackey, Eileen A.; "The Use of Audio-Visual Aids in Training", Social Casework, 21 (9), 1950, pp. 366-371.

Blackey, Eileen; "Selection and Preparation of Faculty for Schools of Social Work", Journal of Education for Social Work, 1 (1), 1965, pp. 5-12.

Block, Alvan M.; "The Dilemma of Social Work Education: Restructuring the Curriculum", Journal of Education for Social Work, 8 (1), 1972, pp. 19-23.

Bloom, B. S.; "Testing Cognitive Ability and Achievement", in N. L. Gage (ed.), Handbook of Research on Teaching (Chicago: Rand McNally, 1963), Ch. 8.

Boehm, Werner W.; "Continuum in Education for Social Work", in E. J. Mullen, J. R. Dumpson, and Associates, Evaluation of Social Intervention (London: Jossey-Baas, 1972), pp. 210-239.

Borman, W. C. and W. R. Vallon; "A View of What Can Happen When Behavioral Expectation Scales are Developed in One Setting and Used in Another", Journal of Applied Psychology, 59, 1974, pp. 197-201.

Boruch, R. F.; "Bibliography: Illustrative Randomized Field Experiments for Program Planning and Evaluation", Evaluation, 2 (1), 1974, pp. 83-87.

Brawley, E. A.; "A Systems Approach to the Evaluation of a Community College Program in Mental Health Work", in Approaches to Innovation in Social Work Education (New York: Council on Social Work Education, 1974), pp. 15-25.

Breedlove, James L.; "Theory Development as a Task for the Evaluator", in E. J. Mullen, J. R. Dumpson and Associates, Evaluation of Social Intervention (London: Jossey-Bass, 1972), pp. 55-70.

Brennan, Clifford E. and Morton L. Arkava; "Students View the Undergraduate Curriculum", Journal of Education for Social Work, 10 (3), 1974, pp. 9-16.

Brenner, M. N. and W. H. Koch; "Continuing Education Among Social Workers: Patterns and Profiles" in Approaches to Innovation in Social Work Education, Council on Social Work Education, New York, 1974, pp. 26-37.

Briar, Scott; "The Age of Accountability", Editoral Page, Social Work, 18, (January 1973), p. 2.

Brigham, T. M.; "Philosophical Issues and Trends in Student Selection Today", Journal of Education for Social Work, 4 (2), 1968, pp. 27-34.

British Association of Social Workers; "Policy Statement on Basic Professional Education for Social Work", Social Work Today, 2 (23), 1972, pp. 16-18.

Brodbeck, May; "Logic and Scientific Method in Research on Teaching", in N. L. Gage (ed.), Handbook of Research on Teaching (Chicago: Rand McNally, 1963), pp. 44-93.

Brown, Malcolm, J; "Social Work Values in a Developing Country", Social Work, 15 (1), 1970, pp. 107-112.

Burns, Mary E.; "Paths to Knowledge: Some Prospects and Problems", Journal of Education for Social Work, 1 (1), 1965, pp. 13-17.

Burnsaka, R. F. and T. D. Hollmann; "An Empirical Comparison of the Relative Effects of Rater Response Biases on Three Rating Scale Formats", Journal of Applied Psychology, 59 (3), 1974, pp. 307-312.

Campbell, D. T.; "From Description to Experimentation: Interpretating Trends as Quasi-Experiments" in C. W. Harris (ed.); Problems in Measuring Change (Madison: University of Wisconsin Press, 1963).

Campbell, Donald R. and Robert F. Borush; "Making the Case for Randomized Assignment to Treatments by Considering the Alternatives: Six Ways in Which Quasi-Experimental Evaluations in Compensatory Education Tend to Underestimate Effects", in C. A. Bennett and A. A. Lumsdaine (eds.); Evaluation and Experiment: Some Critical Issues in Assessing Programs (New York: Academic Press, 1975), pp. 195-275.

Cambpell, J. P., M. D. Dunnette, R. D. Avrey, and L. V. Hellerwik; "The Development and Evaluation of Behaviorally Based Rating Scales", Journal of Applied Psychology, 57, 1973, pp. 15-22.

Cresey, Cannon; "Social Workers: Training and Professionalism" in Trevor Pateman, (ed.); Counter Course: A Handbook for Course Criticism (Harmondsworth, Middlesex: Penguin, 1972), pp. 247-263.

Carlson, Raymond W.; "Expanding Educational Assumptions for Social Work Education", Journal of Education for Social Work, 10 (3), 1974, pp. 17-24.

Carrier, John and Ian Kendall; "Social Policy and Social Change--Explanations of the Development of Social Policy", Journal of Social Policy, 2 (3), 1973, pp. 209-224.

Centra, J. A.; "Two Studies on the Utility of Student Ratings for Improving Teaching", SIR Report No. 2, Princeton, N.J.: Educational Testing Service, 1972.

Centra, J. A.; "Self Rating of College Teachers: A Comparison With Student Ratings", Journal of Educational Measurement, 10, 1973, pp. 287-295.

Centra, J. A.; "The Relationship Between Student and Alumni Ratings of Teachers", Research Bulletin 73-79, Princeton, N.J.: Educational Testing Service, 1973.

Centra, J.; "College Teaching: Who Should Evaluate It?" Findings, 1 (1), Educational Testing Service, Princeton, N.J., 1974.

Centra, John A. and R. L. Linn; "Student Points of View in Ratings of College Instruction", Educational and Psychological Measurement, 36, 1976, pp. 693-703.

Central Council for Education and Training in Social Work; "Social Work: Setting the Course for Social Work Education", CCETSW Report 1, London, 1971-1973.

Central Council for Education and Training in Social Work; "Education and Training in Social Work", Discussion Paper No. 10, London, November, 1975.

Central Council for Education and Training in Social Work; "Social Work: The Pattern of Professional Training", CCEWSW Leaflet 2, London, undated.

Cherryholmes, C. H.; "Some Current Research on Effectiveness of Educational Simulations", Review of Educational Research, 38, (1968), pp. 160-176.

Chommie, Peter W. and Joe Hudson; "Evaluation of Outcome and Process", Social Work, November 1974, pp. 682-687.

Chronbach, L. J.; "Evaluation for Course Improvement", Teachers College Record, 64, 1963, pp. 672-683.

Chronbach, Lee J.; "Validation of Educational Measurements", in Proceedings of the 1969 Invitational Conference on Testing Problems (Princeton, N.J.: Educational Testing Service, 1969), pp. 35-52.

Chronbach, Lee J. and Paul Meehl; "Construct Validity in Psychological Tests", Psychological Bulletin, 52, 1955, pp. 281-302.

Clemenger, Florence; "Assessment of Students' Performance in the Classroom", in Marguerite Pohek (ed.), Teaching and Learning in Social Work Education (New York: Council on Social Work Education, 1970), pp. 27-33.

Cohen, David K. and Michael S. Garet; "Reforming Educational Policy with Applied Social Research", Harvard Educational Review, 45 (1), 1975, pp. 17-41.

Coleman, J. S.; "Evaluating Educational Programs: A Symposium", The Urban Review, 3 (4), 1969, pp. 6-8.

Costin, Lela B.; "Values in Social Work Education: A Study", Social Service Review, 38 (3), 1964, pp. 271-280.

Coughlin, Bernard J.; "Reconceptualizing the Theoretical Base of Social Work Practice", in Lillian Ripple (ed.), Innovations in Teaching Social Work Practice (New York: Council on Social Work Education, 1970), pp. 58-71.

Coyle, Grace L.; "The Role of the Teacher in the Creation of an Integrated Curriculum", Social Work Journal, 33 (2), 1955, pp. 73-82.

Crittendon, K. S., and J. L. Norr; "Student Values and Teacher Evaluation", Sociometry, 36, 1974, pp. 143-151.

Davies, Martin; "The Current Status of Social Work Research", British Journal of Social Work, 4 (3), 1974, pp. 282-303.

Davies, Martin, and Elinor Kelly; "The Social Worker, the Client and the Social Anthropologist", British Journal of Social Work, 6 (2), 1976, pp. 214-231.

Doyle, K. O. Jr., I. I. Bejar, and S. E. Whitely; "Norms for the Student Opinion Survey", Measurement Services Center, Minneapolis, University of Minnesota, 1974.

Doyle, K. O. and S. E. Whitely; "Student Ratings as Criteria for Effective Teaching", American Educational Research Journal, 11 (3), 1974, pp. 259-274.

Dreeben, Robert; "The School as a Workplace", in R. M. W. Travers (ed.), Second Handbook of Research on Teaching, American Educational Research Association (Chicago: Rand McNally, 1973), pp. 450-473.

Drucker, A. T. and H. H. Remmers; "Do Alumni and Students Differ in Their Attitudes Toward Instructors?", Journal of Educational Psychology, 42, 1951, pp. 129-143.

Duehn, Wayne D. and Enola K. Proctor; "A Study of Cognitive Complexity in the Education for Social Work Practice", Journal of Education for Social Work, 10 (2), 1974, pp. 20-26.

Ebel, Robert; "Must All Tests Be Valid?" American Psychologist, 16, 1961, pp. 640-647.

Edwards, Ward and Marcia Guttentag; "Experiments and Evaluations: A Re-examination", in C. A. Bennett and A. A. Lumsdaine (eds.), Evaluation and Experiment: Some Critical Issues in Assessing Programs (New York: Academic Press, 1975), pp. 409-463.

Eisner, E. W.; "Educational Objectives: Help or Hinderance?" School Review, 75, 1967, pp. 250-260.

Entwistle, N. J.; "Complementary Paradigms for Research and Development Work in Higher Education", in W. A. Verrek (ed.), Methodological Problems in Research and Development in Higher Education (Amsterdam: Swets & Zeitlinger, 1974), pp. 75-88.

Feldman, Ronald; "Toward the Evaluation of Teaching Competence in Social Work", Journal of Education for Social Work, 8 (2), 1972, pp. 5-15.

Feldstein, Donald L.; "Heliocentric Perspective on Social Work Education", in E. J. Mullen, J. R. Dumpson, and Associates, Evaluation of Social Intervention, (London: Jossey-Bass, 1972), pp. 240-250.

Fellin, Phillip, A.; "Long-range Planning Implications for Social Work Education", Journal of Education for Social Work, 10 (2), 1974, pp. 27-33.

Ferman, Louis A.; "Some Perspectives on Evaluating Social Welfare Programs", Annals of the American Academy of Arts and Sciences, 385, 1969, pp. 143-156.

Finestone, Samuel; "A Memorandum on Research Into Social Work Education", Journal of Education for Social Work, 1 (1), 1965, pp. 18-25.

Flanigan, Beverly; "Planned Change and Contract Negotiation as an Instructional Model", Journal of Education for Social Work, 10 (2), 1974, pp. 34-39.

Forbes, K. M.; "Training and Function", British Journal of Psychiatric Social Work, 2 (5), 1951.

Freeman, H. E.; "Outcome Measures and Social Action: An Immodest Proposal for Redirecting Research Efforts", American Sociologist, 7, 1972, pp. 26-30.

Freund, Paul A.; "Ethical Aspects of Experimentation on Human Subjects", Daedalus, 98 (2), 1969, pp. vii-xiv.

Frey, Louise A.; "The Evaluation of Teacher Competence in Continuing Education", Social Work Education Reporter, 20 (3), 1972, pp. 43-47.

Gage, N. L.; "The Appraisal of College Teaching", Journal of Higher Education, 32, 1961, pp. 17-22.

Gage, N. L.; "Paradigms for Research on Teaching", in N. L. Gage (ed.), Handbook of Research on Teaching (Chicago: Rand McNally, 1963), pp. 94-141.

Gage, N. L.; "Successful Teacher Behavior: Can Science Contribute to the Art of Teaching", in Marguerite Pohek (ed.), Teaching and Learning in Social Work Education (New York: Council on Social Work Education, 1970), pp. 1-12.

Gagné, F. and D. Allaire; "Summary of Research Data on the Reliability and Validity of a Measure of Dissatisfaction Derived from Reality-Desires Discrepancies", Institute National de la Recherche Scientifique, Universite du Quebec, Quebec, 1974.

Geismar, Ludwig L.; "Some Problems in Practice-oriented Research", Social Work, 6 (4), 1961, pp. 50-56.

Geismar, L.; "Thirteen Evaluative Studies", in E. J. Mullen, J. R. Dumpson, and Associates, Evaluation of Social Intervention (London: Jossey-Bass, 1972), pp. 15-38.

Getzels, J. W. and E. G. Guba; "Role, Role Conflict, and Effectiveness", American Sociological Review, 19, 1954, pp. 164-175.

Gilbert, J.P., R. L. Light and F. Mosteller; "Assessing Social Innovations: An Empirical Base for Policy", in C. A. Bennett and A. A. Lumsdaine (eds.), Evaluation and Experiment: Some Critical Issues in Assessing Social Programs (New York: Academic Press, 1975), pp. 39-182.

Glaser, R. and W. W. Coley; "Instrumentation for Teaching and Instructional Management", in R. M. W. Travers (ed.), Second Handbook of Research on Teaching (Chicago: Rand McNally, 1973), pp. 832-857.

Goldstein, Harris K.; "The Measurement of Attitudes and Research Knowledge", Tulane Studies in Social Welfare, New Orleans, Tulane University, 1968.

Goldstein, Harris K. and Linda Horder; "Suggested Teaching Plans for Maximizing Research Learning of Three Types of Social Work Students", Journal of Social Work Education, 10 (3), 1974, pp. 30-35.

Graham-Smith, H. and J. W. D. Davies; "Professional Training and Afterwards--Two Former Students Reflect", British Journal of Psychiatric Social Work, 11 (1), 1964, pp. 3-8.

Grant, C. W.; "Faculty Allocation of Effort and Student Course Evaluations", Journal of Educational Research, 64, 1971, pp. 405-410.

Green, Judy Kopp and William R. Morrow; "Precision Social Work: General Model and Illustrative Student Projects With Clients", Journal of Education for Social Work, 8 (3), 1972, pp. 19-29.

Grey, A. N. and H. E. Dermody; "Reports of Casework Failure", Social Casework, November 1972, pp. 534-543.

Guba, Egon G. and Daniel L. Stufflebeam; "Evaluation: The Process of Stimulating, Aiding, and Abetting Insightful Action", address delivered at Second National Symposium for Professors of Educational Research, Nov. 21, 1968, Columbus, Ohio, Evaluation Centre, College of Education, Ohio State University.

Guba, Egon G.; "Problems in Utilizing the Results of Evaluation", Journal of Research and Development in Education, 8 (3), 1975, pp. 42-54.

Gurin, Arnold; "Education for Changing Practice", in Alfred J. Kadushin (ed.), Shaping the New Social Work (New York: Columbia University Press, 1973), pp. 169-198.

Guttentag, M.; "Models and Methods in Evaluation Research", Journal of Theory of Social Research, 1 (1), 1971, pp. 75-95.

Guzetta, C.; "Concepts and Precepts in Social Work Education", Journal of Education for Social Work, 2 (2), 1966, pp. 40-47.

Guzetta, C.; "Curriculum Alternatives", Journal of Education for Social Work, 8 (2), 1972, p. 25.

Hack, Kenneth A.; "A Study of Predictors of Success on a Social Work Course", British Journal of Social Work, 3 (2), 1973, pp. 189-196.

Hale, M. P.; "Focus and Scope of a School of Social Work", Journal of Education for Social Work, 3 (2), 1967.

Havelock, R. G.; "The Utilization of Educational Research and Development", British Journal of Educational Technology, 2 (2), 1971, pp. 84-98.

Hayes, Dorothy D. and Barbara K. Varley, "The Impact of Social Work Education on Students' Values", Social Work, 19 (3), July 1965, pp. 4-46.

Hearn, Gordon; "Accountability in Social Work Education", in Boyd E. Oviatt (ed.), Evaluation and Accountability in Social Work Education (Salt Lake City: Graduate School of Social Work, University of Utah, 1973), pp. 4-15.

Heynes, R. W. and R. Lippitt; "Systematic Observational Techniques", in G. Lindzey (ed.), Handbook of Social Psychology (Cambridge: Mass.: Addison-Wesley, 1954), pp. 370-404.

Hilton, Elizabeth T.; "Field Trial Designs in Gauging the Impact of Fertility Planning Programs", in C. A. Bennett and A. A. Lumsdaine (eds.), Evaluation and Experiment: Some Critical Issues in Assessing Social Programs (New York: Academic Press, 1975), pp. 319-408.

Holden, F. J.; "A New Look at Instructional Technology", Canadian Training Methods, May-June, 1972, pp. 15-18.

Holland, T., L. Olsen and P. Chatterjee; "Evaluation Research: Some Possible Contexts of Theory Failure", Journal of Sociology and Social Welfare, 3 (4), 1976, pp. 384-402.

Hollis, Florence; "And What Shall We Teach? The Social Work Educator and Knowledge", Social Service Review, 42 (2), 1968, pp. 184-196.

Hoshino, G.; "Social Services, The Problem of Accountability", Social Service Review, 47, 1973, pp. 373-383.

Howery, Victor I." "Continuing Education: Program Development, Administration, and Financing", Journal of Education for Social Work, Winter, 1974, pp. 34-41.

Hoyt, D. P.; "Test Reliability Obtained by Analysis of Variance", Psychometrika, 6, 1941, pp. 153-160.

Hoyt, D. P.; "Measurement of Instructional Effectiveness", Research in Higher Education, 1, 1973, pp. 367-378.

Ingleby, David; "Ideology and the Human Sciences: Some Comments on the Role of Reification in Psychology and Psychiatry" in Trevor Pateman (ed.), Counter Course: A Handbook for Course Criticism (Harmondsworth: Penguin, 1972), pp. 51-81.

Isaacson, R. L. et al.; "Dimensions of Student Evaluations of Teaching", Journal of Educational Psychology, 55 (6), 1964, pp. 344-351.

Jenkins, David; "Integrated Studies Project" in Schools Council, Evaluation in Curriculum Development: Twelve Case Studies (London: Macmillan, 1973), pp. 70-79.

Jonas, Hans; "Philosophical Reflections on Experimenting With Human Subjects", Daedalus, 98 (2), 1969, pp. 219-247.

Jones, D. and Huczynski, A.; "DHSS GMD Course Evaluation Study, Report of the Evaluation Stage", School of Management Studies, Polytechnic of Central London, London, 1974 (mimeo).

Jones, W. C., and E. J. Borgatta; "Methodology of Evaluation" in E. J. Mullen, J. R. Dumpson, and Associates, Evaluation of Social Intervention (London: Jossey-Bass, 1972), pp. 39-54.

Kadushin, Alfred, "Interview Observation as a Teaching Device", Social Casework, XXXVII (7), 1956, pp. 334-341.

Kadushin, Alfred; "Two Problems of the Graduate Program: Level and Content", Journal of Education for Social Work, 1 (1), 1965, pp. 33-46.

Kadushin, Alfred and George Kelling; "Final Report, An Innovative Program in Social Work Education: The 3-2 Program", Madison, University of Wisconsin, mimeographed.

Kadushin, Alfred and George Kelling; "An Innovative Program Reducing Length of Training: Evaluation Procedures and Outcome", Journal of Education for Social Work, 13 (2), 1977, pp. 68-75.

Kahn, S. B. and J. Weiss; "The Teaching of Affective Responses" in R. M. W. Travers (ed.), Second Handbook of Research on Teaching (Chicago: Rand McNally, 1973), pp. 759-804.

Katz, Daniel; "Feedback in Social Systems: Operational and Systemic Research on Production, Maintenance, Control, and Adaptive Functions", in C. A. Bennett and A. A. Lumsdaine (eds.), Evaluation and Experiment: Some Critical Issues in Assessing Social Programs (New York: Academic Press, 1975), pp. 465-516.

Kendall, P.; "Evaluating an Experimental Programme in Medical Education" in M. B. Miles (ed.), Innovation in Education (New York: Teachers' College Press, Columbia University, 1964).

Kidneigh, John C. and Horace W. Lundberg; "Are Social Work Students Different?", Social Work, 3 (2), 1958, pp. 57-61.

Kiresuk, T. J. and R. E. Sherman; "Goal-Attainment Scaling: A General Method for Evaluating Comprehensive Community Mental Health Programs", Community Mental Health Journal, 4, 1968, pp. 443-453.

Kogan, N. "Attitudes Toward Old People: the Development of a Scale and an Examination of Correlates", Journal of Abnormal Psychology, 62, 1961, pp. 616-622.

Kogan, Leonard S. and Ann W. Shyne; "Tender-Minded and Tough-Minded Approaches in Evaluative Research", Welfare in Review, 4 (2), 1966.

Kolevson, Michael F.; "Integrational Teaching Modalities in Social Work Education: Promise or Pretence?", Journal of Education for Social Work, 11 (2), 1975, pp. 60-67.

Kolevson, M. S. and K. T. Wiltze; "Student Ratings and Teacher Effectiveness: A Reappraisal", Journal of Education for Social Work, 9 (2), 1973.

Knowles, Malcolm S.; "Innovations in Teaching Styles and Approaches Based Upon Adult Learning", Journal of Education for Social Work, 8 (2), 1972, pp. 32-39.

Kuder, G. F. and M. W. Richardson; "The Theory of the Estimation of Test Reliability", Psychometrika, 2, 1937, pp. 151-160.

Lauffer, Armand; "Continuing Education as Problem-Focused Extension", Journal of Education for Social Work, 8 (3), 1972, pp. 40-49.

Leonard, P.; "Experiments in Post-Professional Education", Social Work Today, 1 (4), 1970, p. 5.

Leonard, Peter; "Explanation and Education in Social Work", British Journal of Social Work, 5 (3), 1975, pp. 325-333.

Levine, Abraham; "Evaluating Program Effectiveness and Efficiency", Welfare in Review, 5 (2), 1967, pp. 1-11.

Levine, Abraham S.; "Cost-Benefit Analysis and Social Welfare Program Evaluation", Social Service Review, 42 (2), 1968, pp. 173-183.

Levine, Harold and Christine McQuire; "Role Playing as an Evaluative Technique", Journal of Educational Measurement, 5 (1), 1968, pp. 1-8.

Levinthall, C. F., L. M. Lansky and O. E. Andrews; "Student Evaluations of Teacher Behaviors as Estimations of Real-ideal Discrepancies: a Critique of Teacher Rating Methods", Journal of Educational Psychology, 62, 1971, pp. 104-109.

Levy, Charles S.; "A Framework for Planning and Evaluating Social Work Education", Journal of Education for Social Work, 8 (2), 1972, pp. 40-47.

Lewis, John and Faith Gibson; "The Teaching of Some Social Work Skills: Toward a Skills Laboratory", British Journal of Social Work, 7 (2), 1977, pp. 189-209.

Lide, Pauline D.; "An Experiment in Teaching Social Casework", Social Service Review, XXXIII (3), 1959, pp. 253-259.

Lifshitz, M.; "Quality Professionals: Does Training Make a Difference A Personal Construct Theory Study of the Issue", British Journal of Clinical Psychology, 13, 1974, pp. 183-189.

Light, R. J. and P. V. Smith; "Choosing a Future: Strategies for Designing and Evaluating New Programs", Harvard Educational Review, 40, 1970, pp. 1-28.

Lindsay, Richard P.; "Comments" in Boyd E. Oviatt (ed.), Evaluation and Accountability in Social Work Education (Salt Lake City; Graduate School of Social Work, University of Utah, 1973), pp. 15-21.

Linn, M. W., S. Stein, J. S. Carter, and B. S. Linn; "The C.A.R.E. Project: Interdisciplinary and Community-Oriented Approach to Training Social Work Students", Journal of Education for Social Work, 10 (2), 1974, pp. 47-50.

Litwak, E.; "Comments on Professor Trow's Paper: Towards Multimodel Theory of Evaluation", in M. C. Wittrock and D. E. Wiley (eds.), The Evaluation of Instruction: Issues and Problems (New York: Holt, Rinehart & Winston, 1970), pp. 305-317.

Lodge, Richard, "Towards Accountability and Evaluation in Social Work Education" in Boyd E. Oviatt (ed.), Evaluation and Accountability in Social Work Education (Salt Lake City: Graduate School of Social Work, University of Utah, 1973), pp. 23-26.

Loeb, Peter C. and John A. Slosar Jr.; "Sociatrogenic Dysfunctions: A Concern for Social Work Education", Journal of Education for Social Work, 10 (2), 1974, pp. 51-58.

Lovell, G. D. and C. F. Haner; "Forced-choice Scales Applied to College Faculty Rating", Educational and Psychological Measurement, 15, 1955, pp. 291-304.

Lundberg, Horace; "Comments" in Boyde E. Oviatt (ed.), Evaluation and Accountability in Social Work Education (Salt Lake City: Graduate School of Social Work, University of Utah, 1973), pp. 34-38.

Maas, Henry S.; "Purpose, Participation, and Education" in Lillian Ripple, (ed.), Innovations in Teaching Social Work Practice (New York: Council on Social Work Education, 1970), pp. 9-21.

MacLauchlin, A. M. and H. Pincus; "The Role of a Clinical Social Worker in a Changing Society", Journal of Education for Social Work, 10 (2), 1974, pp. 59-64.

Mager, R. F. and C. Clark; "Explorations in Student-Controlled Instruction", Psychological Reports, 13, 1963, pp. 71-76.

Mann, John; "Technical and Social Difficulties in the Conduct of Evaluative Research" in Francis G. Caro (ed.), Readings in Evaluative Research (New York: Russell Sage Foundation, 1971), pp. 175-184.

Martin, O. A.; "Assessing Training Effectiveness", Department of Employment and Productivity, London, October 1968 (mimeo).

Maslow, A. H. and W. Zimmerman; "College Teaching Ability, Scholarly Activity and Personality", Journal of Educational Psychology, 47, 1956, pp. 185-189.

McKeachie, W. J.; "Student Ratings of Faculty", AAUP Bulletin, 55, 1969, pp. 439-444.

McLean Ann L. and Joan F. Robertson; "Systematic Integration of Consumer Input into Curriculum Revision", Journal of Education for Social Work, 11 (2), 1975, pp. 84-90.

McQuire, Christine H. and David Babbott; "Simulation Technique in the Measurement of Problem Solving Skills", Journal of Educational Measurement, 4, 1967, pp. 1-10.

Mead, Margaret; "Research with Human Beings: A Model Derived from Anthropological Field Practice", Daedalus, 98 (2), Spring 1969, pp. 361-386.

Medley, D. M., and H. E. Mitzel; "Measuring Classroom Behavior by Systematic Observation" in N. L. Gage (ed.), Handbook of Research on Teaching (Chicago: Rand McNally, 1973), pp. 247-328.

Meinert, Roland G.; "Simulation Technology: A Potential Tool for Social Work Education", Journal of Education for Social Work, 8 (3), 1972, pp. 50-59.

Miller, C. and M. Parlett; "Up to the Mark--a Research Report on Assessment", Occasional Paper 13, Centre for Research in the Educational Sciences, University of Edinburgh, January 1973.

Mullen, Edward J.; "Evaluative Research on the Effects of Professional Intervention" in Paul C. Vrooman (ed.), Trans-Disciplinary Issues in Social Welfare (Waterloo, Canada: Graduate SChool of Social Work, Waterloo Lutheran University, 1972), pp. 17-35.

Munday, B.; "What is Happening to Social Work Students?", Social Work Today, 3 (6), 1972, p. 3.

Neely, M.; "A Teacher's View of Teacher Evaluation", Improving College and University Teaching, 16, 1968, pp. 207-209.

Papell, Catherine; "Teaching the Technology of Social Work: Cumulative Methods to Unitary Whole" in Kay Dea (ed.), New Ways of Teaching Social Work Practice (New York: Council on Social Work Education, 1972), pp. 1-18.

Parlett, Malcolm; "Evaluating Innovations in Teaching" in H. J. Butcher and E. Rudd (eds.), Contemporary Problems in Higher Education (London: McGraw-Hill, 1972), pp. 144-153.

Parlett, Malcolm and David Hamilton; "Evaluation as Illumination, A New Approach to the Study of Innovatory Programs", Occasional Paper, Centre for Research in the Educational Sciences, University of Edinburgh, 1972.

Parlett, Malcolm and David Hamilton; "Evaluation as Illumination: A New Approach to the Study of Innovatory Programs" in Gene V. Glass (ed.), Evaluation Studies Review Annual, Vol. 1 (Beverly Hills: Sage Publications, 1976), pp. 140-157.

Parsons, Talcott; "Research With Human Subjects and the Professional Complex", Daedalus, 98 (2), 1969, pp. 325-360.

Patton, Michael Quinn; "Stages and Strategies of Evaluation Research", paper prepared for Evaluation Session, American Montessori Society Convention, Granby, Colorado, June, 1975.

Payne, Paul; "Use of a Situation Test for Predicting Counselor Performance", Journal of Counseling Psychology, 15 (6), 1968, pp. 512-516.

Perlman, Helen Harris; "Teaching Casework by the Discussion Method", Social Service Review XXIV, (3), 1950, pp. 334-346.

Perlman, Helen Harris; "The Lecture as a Method in Teaching Casework", Social Service Review, XXV, (1), 1951, pp. 19-32.

Piper, David Warren; "Middlemuch: Staff Development in Universities" in W. A. Verreck (ed.), Methodological Problems in Research and Development in Higher Education (Amsterdam: Swets & Zeitlinger, 1974), pp. 143-160.

Posner, Michael and S. W. Keele; "Skill Learning" in R. M. W. Travers (ed.), Second Handbook of Research on Teaching (Chicago: Rand McNally, 1973), pp. 805-831.

Puckett, T. C.; "Social Science Methods and Social Work Research . . .", Journal of Applied Social Studies, 3, 1971, pp. 29-38.

Rayder, N. F.; "College Student Ratings of Instructors", Journal of Experimental Education, 37 (2), 1968, pp. 76-81.

Rees, R. D.; "Dimensions of Students' Points of View in Rating College Teachers", Journal of Educational Psychology, 60, 1969, pp. 476-482.

Robinson, James; "Experimental Research in Social Casework", British Journal of Social Work, 1 (4), 1971, pp. 463-479.

Rodin, M. and B. Rodin; "Student Evaluations of Teachers", Science, 177, 1972, pp. 1164-1166.

Rosen, A. and Dineh Lieberman; "Experimental Evaluation of Interview Performance of Social Work", Social Service Review, 46 (3), 1972, pp. 395-412.

Rosenshine, B. and N. Furst; "The Use of Direct Observation to Study Teaching", in R. M. W. Travers (ed.), Second Handbook of Research on Teaching (Chicago: Rand McNally, 1973), pp. 122-183.

Ross, C.; "An Educational Philosophic Inventory: An Instrument for Measuring Change and Determining Philosophic Perspective", The Journal of Educational Thought, 4 (1), 1970, pp. 20-26.

Rossi, P.; "Evaluating Educational Programs: A Symposium", The Urban Review, 3 (4), 1969, pp. 17-18.

Rossi, Peter H.; "Practice, Method and Theory in Evaluating Social Action Programs", in James L. Sundquist (ed.), On Fighting Poverty (New York: Basic Books, 1969).

Rothman, B. and J. L. Vigilante; "Curriculum Plannning in Social Work Education", Journal of Education for Social Work, 10 (2), 1974, pp. 76-85.

Ryle, A. and D. Breen; "Change in the Course of Social Work Training:A Repertory Grid Study", British Journal of Medical Psychology, 47, 1974, pp. 139-148.

Ryle, A. and D. Breen; "Social Work Tutors' Judgements of Their Students", British Journal of Medical Psychology, 47, 1974, pp. 149-152.

Sales, Esther and Elizabeth Navarre; "Teaching in Groups, Individual and Group Supervision in Field Instruction: A Research Report", School of Social Work, University of Michigan, 1970.

Sanders, Daniel S.; "Educating Social Workers for the Role of Effective Change Agents in a Multicultural, Pleuralistic Society", Journal of Education for Social Work, 10 (2), 1974, pp. 86-91.

Sax, Gilbert; "The Use of Standardized Tests in Evaluation", in W. James Popham (ed.), Evaluation in Education: Current Applications (Berkeley: McCutchan, 1974), pp. 245-308.

Scriven, Michael; "The Methodology of Evaluation" in R. E. Stake (ed.), AERA Monograph Series in Education, No. 1 (Chicago: Rand McNally, 1967), pp. 39-89.

Scriven, M.; "Evaluating Educational Programs: A Symposium", The Urban Review, 3 (4), 1969, pp. 20-22.

Scriven, Michael; "Evaluation Perspectives and Procedures", in W. James Popham (ed.), Evaluation in Education: Current Applications (Berkeley: McCutchan, 1974), pp. 7-33.

Scriven, Michael A.; "Prose and Cons About Goal-Free Evaluation", in W. James Popham (ed.), Evaluation in Education: Current Applications (Berkeley: McCutchan, 1974), pp. 34-67.

Segal, Brian and Michael J. Austin: "Innovations in Field Instruction: Community College Teaching and Staff Development", Journal of Education for Social Work , Winter, 1974, pp. 77-80.

Selby, Lola; "The University of California Generalist Model for Teaching Social Work Methods" in Kay Dea (ed.), New Ways of Teaching Social Work Practice (New York: Council on Social Work Education, 1972), pp. 19-32.

Sharon, A. T.; "Eliminating Bias from Student Ratings of College Instructors", Journal of Applied Psychology, 54, 1970, pp. 278-281.

Sharon, A. T. and C. J. Bartlett; "Effect of Instructional Conditions in Producing Leniency on Two Types of Rating Scales", Personnel Psychology, 22, 1969, pp. 251-263.

Sharwell, George S.; "Can Values be Taught? A Study of Two Variables Related to Orientation of Social Work Graduate Students Toward Public Dependency", Journal of Education for Social Work, 10 (2), 1974, pp. 99-105.

Sheldrake, Peter; "Uncertain Prognosis: Behavioral Sciences in Undergraduate Medicine", Occasional Paper No. 17, Centre for Research in the Educational Sciences, University of Edinburgh, 1973.

Sheldrake, P.; "The Politics of the Curriculum" in W. A. Verreck (ed.), Methodological Problems in Research and Development in Higher Education (Amsterdam: Swets & Zeitlinger, 1974), pp. 89-104.

Sherwood, Clarence C.; "Issues in Measuring the Results of Action Programs", Welfare in Review, 5 (7), 1967, pp. 13-18.

Sinfield, Adrian; "Which Way for Social Work", Fabian Tract 393, Fabian Society, 1969.

Skidmore, Rex A.; "Evaluation and Accountability in Schools of Social Work", in Boyd E. Oviatt (ed.), Evaluation and Accountability in Social Work Education (Salt Lake City: Graduate School of Social Work, University of Utah, 1973), pp. 27-34.

Smith, B. O.; "Teaching and Testing Values", in Proceedings of the 1965 Invitational Conference on Testing Problems (Princeton, N.J.: Educational Testing Service, 1960), pp. 50-59.

Smith, G.; "Staff Tensions in Social Work Education", Social Work Today, 3 (6), 1972, p. 8.

Smith, P. C. and J. G. Kendall; "Retranslation of Expectations: An Approach to the Construction of Unambiguous Anchoring of Rating Scales", Journal of Applied Psychology, 47, 1963, pp. 149-155.

Soloman, D., L. Rosenberg, and W. E. Benzdek; "Teacher Behavior and Student Learning", Journal of Educational Psychology, 55, 1964, pp. 23-30.

Somers, Mary Louise; "Contributions of Learning and Teaching Theories to the Explication of the Role of the Teacher in Social Work Education", Journal of Education for Social Work, 5 (2), 1969, pp. 61-73.

Somers, Mary Louise; "Dimensions and Dynamics of Engaging the Learner", Journal of Education for Social Work, 7 (3), 1971, pp. 49-57.

Specht, Harry; "Issues and Problems in Utilizing a Unitary Method" in H. Specht and A. Vickery (eds.), Integrating Social Work Methods (London: Allen & Unwin, 1977), pp. 248-255.

Stake, R. E.; "Toward a Technology for the Evaluation of Educational Programs", AERA Monograph Series on Curriculum Evaluation (Chicago: Rand McNally, 1967).

Stein, Herman D.; "Cross-Currents in Practice, Undergraduate and Graduate Education in Social Work", Journal of Education for Social Work, 1 (1), 1965, pp. 56-67.

Stein, Herman D.; "Reflections on Competence and Ideology in Social Work Education", Journal of Education for Social Work 5 (1), 1969, pp. 81-90.

Stein, Shayna, Margaret W. Linn and James Furdon; "Predicting Social Work Student Performance", Journal of Education for Social Work, 10 (3), 1974, pp. 85-92.

Stern, G. G.; "Measuring Noncognitive Variables in Research on Teaching" in N. L. Gage (ed.), Handbook of Research on Teaching (Chicago: Rand McNally, 1963), pp. 398-434.

Stevenson, O.; "Problems in the Use of Theory in Social Work Education", British Journal of Psychiatric Social Work, 9 (1), 1967, pp. 27-29.

Stevenson, Olive; "Knowledge for Social Work", British Journal of Social Work, 1 (2), 1971, pp. 225-237.

Stufflebeam, Daniel L.; "Alternative Approaches to Educational Evaluation: A Self-Study Guide for Educators" in W. James Popham, Evaluation in Education: Current Applications (Berkeley: McCutchan, 1974), pp. 97-143.

Stumpf, Jack; "Developing the Social Work Practice Curriculum at San Diego State College: A Further Statement", in Kay Dea (ed.), New Ways of Teaching Social Work Practice (New York: Council on Social Work Education, 1972), pp. 33-47.

Suchman, E.; "Evaluating Educational Programs: A Symposium", The Urban Review, 3 (4), 1969, pp. 15-17.

Taber, M. and I. Shapiro; "Social Work and Its Knowledge Base: A Content Analysis of the Periodical Literature", Social Work, 10 (4), 1965, pp. 100-106.

Taber, M.; "A Sampling of Techniques for Evaluation Research in Field Instruction", Social Work Education Reporter, 15 (3), 1967, pp. 22-25.

Tallmadge, G. K. and J. W. Shearer; "Relationships Among Learning Styles, Instructional Methods, and the Nature of Learning Experiences", Journal of Educational Psychology, 60 (3), 1969, 222-230.

Tatsouka, Maurice M. and David V. Tiedeman;"Statistics as an Aspect of Scientific Method in Research on Teaching" in N. L. Gage (ed.), Handbook of Research on Teaching (Chicago: Rand McNally, 1963), pp. 142-170.

Taylor, P. A. and T. O. Maguire; "A Theoretical Evaluation Model", Manitoba Journal of Education Research, 1, 1966, pp. 12-17.

Thomlison, Raymond J. and Frederick W. Seidl; "An Experiment in Curriculum Innovation in Graduate Social Work Education", Journal of Education for Social Work, 10 (3), 1974, pp. 93-98.

Thorndike, R. L.; "Reliability" in E. F. Linquist (ed.), Educational Measurement (Washington: American Council on Education, 1951), Ch. 15.

Timms, Noel; "Theory and Practice in Social Work Education", Case Conference, 5 (7), 1959, pp. 167-173.

Timms, Noel; "Research into Social Work: A Review Article", British Journal of Social Work, 1 (3), 1971, pp. 345-351.

Torre, Elizabeth; "Student Performance in Solving Social Work Problems and Experience Prior to Entering the MSW Program", Journal of Education for Social Work, 10 (2), 1974, pp. 114-117.

Trow, M.; "Methodological Problems in the Evaluation of Innovation" in M. C. Wittrock and D. E. Wiley (eds.), The Evaluation of Instruction: Issues and Problems (New York: Holt, Rinehart and Winston, 1970), pp. 289-305.

Tully, Brian J.; "Personal Construct Theory and Psychological Changes Related to Social Work Training", British Journal of Social Work, 6 (4), 1976, pp. 481-499.

Tyler, R. W.; "Distinctive Attributes of Education for the Professions", Social Work Journal, 33 (2), 1952, pp. 55-62.

Vance, Patricia V.; "The Changing Scene in Social Work Education", Child Welfare, 11 (7), 1972, pp. 413-422.

Vanden Berg, M.; "The Nature of Professional Education", Case Conference, 15 (7), 1968, pp. 263-268.

Varley, Barbara K.; "Socialization in Social Work Education", Social Work, 8 (3), July 1963, pp. 102-109.

Varley, Barbara K.; "Social Work Values: Changes in Value Commitments of Students from Admission to M.S.W. Graduation", Journal of Education for Social Work, 4 (2), 1968, pp. 67-76.

Vickery, Anne; "Use of Unitary Models in Education for Social Work" in H. Specht and A. Vickery (eds.), Integrating Social Work Methods (London: Allen & Unwin, 1977), pp. 233-245.

Vigilante, Joseph L.; "Between Values and Science: Education for the Profession During a Moral Crisis or Is Proof Truth", Journal of Education for Social Work, 10 (3), 1974, pp. 107-115.

Vinter, R. D. and S. Seigel; "Proposal for Schools of Social Work to Establish a National Information System", Ann Arbor, Michigan, 1971.

Walberg, H. J.; "Can Educational Research Contribute to the Practice of Teaching?", Journal of Education for Social Work, 4 (2), 1968.

Walker, Walter L.; "Changing Thought and Action Styles for Students and Faculty: Imperatives for Social Work Education?", Journal of Education for Social Work, 8 (2), 1972, pp. 56-63.

Webb, W. B. and C. Y. Nolan; "Student, Supervisor, and Self-Rating of Instructional Proficiency", Journal of Educational Psychology, 46, 1955, pp. 42-46.

Weiss, Carol H.; "The Politicization of Evaluation Research", Journal of Social Issues, 24 (4), 1970, pp. 57-68.

Weiss, C. H.; "Between the Cup and Lip", Evaluation, 1 (2), 1973, pp. 49-55.

Weiss, Robert S. and Martin Rein; "The Evaluation of Broad-Aim Programs: A Cautionary Case and a Moral", Annals of the American Academy of Arts and Sciences, 385, 1969, pp. 133-142.

Weiss, Robert S. and Martin Rein; "The Evaluation of Broad-Aim Programs: Experimental Design, Its Difficulties, and an Alternative", Administrative Science Quarterly, 15, 1970, pp. 97-109.

Welch, W. W. and H. J. Walberg; "Pretest Effects in Curriculum Evaluation", American Educational Research Journal, 6, 1970, pp. 605-614.

Wholey, Joseph S.; "Planning Useful Evaluations", mimeographed for an Evaluation Workshop, Health Resources Administration, U.S. Department of Health, Education and Welfare, January, 1976.

Wilson, E.; "A Reply to Brian Munday: Another View of What is Happening to Social Work Students", Social Work Today, 3 (9), 1972, p. 15.

Wilson, James; "Evaluation of Learning in Secondary School Mathematics" in B. Bloom, J. T. Hastings and G. F. Madeus, Handbook of Formative and Summative Evaluation of Student Learning (New York: McGraw-Hill, 1971), pp. 687-688.

Withall, J.; "Development of a Technique for the Measurement of Socio-emotional Climate in Classrooms", Journal of Experimental Education, 17 (1949), pp. 347-361.

Withall, J.; "The Development of a Climate Index", Journal of Educational Research, 45 (1951), pp. 93-99.

Witmer, Helen L.; "What is Evaluation?", Casework Papers, New York: Family Service Association, 1954.

Wold, Herman; "Causal Inference from Observational Data: A Review of Ends and Means" in M. C. Wittrock and David E. Wiley, The Evaluation of Instruction: Issues and Problems (New York: Holt Rinehart and Winston, 1970), pp. 351-390.

Wolf, Richard M.; "Data Analysis and Reporting Considerations in Evaluation" in W. James Popham (ed.), Evaluation in Education: Current Applications (Berkeley: McCutchan, 1974), pp. 203-242.

Wolins, Martin; "Measuring the Effects of Social Work Intervention", in Norman A. Polansky (ed.), Social Work Research (Chicago: University of Chicago Press, 1969), pp. 247-272.

Yarbrough, Roy D.; "Operationalism: A Philosophy for Social Work Education", Journal of Education for Social Work, 11 (2), 1975, pp. 116-121.

Yelaja, Shankar A.; "Freedom and Authority in Professional Education", Paper read at the Canadian Association of Professors of Education, Annual Meeting, Canadian Learned Societies, Winnipeg, May 1970 (mimeo).

Younghusband, E.; "Social Work in the Future: Trends and Training" in H. Specht and A. Vickery (eds.), Integrating Social Work Methods (London: Allen & Unwin, 1977), pp. 109-115.

Ziomek, R. L., L. G. Smith, and J. W. Menne; "A Psychometric Analysis of the Ross Educational Philosophical Inventory (REPI)", Educational and Psychological Measurement, 36, 1976, pp. 679-688.

APPENDIX I

INSTRUMENTS FROM STUDY I

Newfargie Cottage,
Gateside,
Fife

10th May, 1976

Dear

Re: NISW Workshop - "A Unitary Approach to Social Work Practice" May 27-28/76

The NISW informs me that you intend to attend the above-mentioned workshop. I imagine that, by this time, you will have received their introductory information package and will be aware that I am attempting to produce an evaluation of the workshop. In that regard, I would very much like to enlist your assistance.

Enclosed you will find a short exercise--the analysis of a case study. I am asking if you will spend a half hour doing it before you attend the workshop and take your 'notes' with you when you go to Chelmsford. I will collect them from you when you arrive. Later, after the workshop experience is over, I will ask you to complete a similar short exercise. I realize that my request is an imposition on your time but it is necessary to the workshop evaluation and is not intended to be very time consuming. Handwritten 'notes' are perfectly acceptable as long as they are legible.

May I take this opportunity to thank you for your anticipated co-operation. It is only with your assistance that an evaluation can be completed and it is my hope that our joint efforts will produce feedback useful to the course tutors.

I look forward to seeing you on May 27th.

Yours sincerely,

COURSE EVALUATOR

Newfargie Cottage
Gateside, Fife
June 12, 1976

Dear

Re: "A Unitary Approach to Social Work Practice"
Jan/76

I know it is beginning to look as if I am going to follow you with questionnaires for the rest of your life but, seriously, this is the last evaluative exercise you will get from me. Enclosed is a case study and instruction sheet. In order to keep my research as "pure" as possible, I cannot explain exactly why I am asking you to spend no more than a half hour making notes on the study. Suffice to say that it is one more effort to evaluate the utility of the workshop.

I realize that this is an imposition on your time and apologize for it. There is, however, no other way of obtaining the necessary information. I must, therefore, ask once more if you would follow the instructions provided (legible handwritten notes are perfectly acceptable) and post your "notes" to me in the enclosed self-addressed envelope.

Thanks again for your co-operation and assistance. The project has become larger than was originally anticipated but my work is going well.

Yours sincerely,

Ken Gordon
Course Evaluator

INSTRUCTIONS TO WORKSHOP PARTICIPANTS

1. Imagine that you are a fieldwork teacher holding a regular seminar group for four C.Q.S.W. students.

2. One student has submitted the attached case study which will be used as the focus for your fourth seminar discussion.

3. For your own benefit, spend no more than one half hour analyzing the case and making notes on the points that you want the discussion to cover.

4. The task in this exercise will focus upon your analysis and suggestions for interventions in the case situation.

5. In order to facilitate later identification please ensure that your name is on the notes. Only the workshop evaluator will be aware of your identity.

CASE STUDY

SURNAME: LUMSDEN

MAN: Reginald, aged 50, occasional employment in manual labour

WOMAN: Elizabeth, aged 41, housewife.

CHILDREN:
Pat (M), aged 18, not residing with parents
John (M), aged 16, not residing with parents
Rose (F), aged 14.
Danny (M), aged 12.
Susan (F), aged 9.

Initial Referral:

This case was originally referred by a health visitor, Miss Anthrope, who had been called in to see the Lumsden children by school authorities who suspected that both Rose and Danny had head lice. Miss Anthrope visited the home, found lice on all three children, and initiated treatment. She also became concerned at the state of the family home and the health of the entire family. She paid a return visit a few days later, and again the next week. On the third home visit Rose was not present and Mrs. Lumsden mentioned that she had not seen Rose for three or four days. Mrs. Lumsden did not, however, appear to be particularly concerned, stating that Rose was always disappearing for a few days at a time but that she always came home sooner or later. Just as the health visitor was preparing to leave, Rose arrived in an obviously intoxicated state. She had been accompanied home by two young men who drove off immediately after "dumping" her at the door. After ensuring that Rose was neither hurt nor ill, Miss Anthrope left. She was, however, very concerned over the 'absolute filth' of the house, Rose's extended absences from home, and Mrs. Lumsden's apparent lack of concern about the situation. She requested that a social worker become involved as she felt that the children were in a state of neglect.

Environmental Circumstances:

A check with the children's school revealed that Rose, Danny and Susan were all very poor attenders and that Mrs. Lumsden either did not care or could do nothing about it. Rose and Danny often arrived for school in a very dirty, dishevelled state and Danny's teacher had gained the impression that he was not very well fed.

It was also found that both Rose and Danny were known at the local youth club. Although Danny is a bit young for the crowd that hangs around there, he is quickly gaining a

reputation as a "tough little guy". On more than one occasion Danny has arrived there with three or four boys when all of them were apparently "high" from sniffing airplane glue. On occasion they have bragged about how intoxicated they have become. One of the centre's volunteers also expressed concern about Rose; suggesting that she was known to drink quite a lot, usually in the company of older boys. She had developed a reputation as an "easy lay" and the volunteer had been given the impression that she sometimes received payment for her favours. Her older brother, John, was well known as a local tough but had not been seen at the centre for some months. He was known to have had a good deal of trouble with the police and it was suggested that he might be incarcerated at this time.

The Lumsden's home is within one of the city's poorer neighbourhoods. The district is well known for its many social problems and poor housing stock.

Home Visit:

A home visit was made on March 16th. Mrs. Lumsden was home with Rose who apparently had been off school, suffering from the 'flu, for a couple of days. She was now feeling much better and thought she would be returning to school shortly. The house is in very run-down condition and is apparently owned by Mrs. Lumsden's elderly mother who is now in a nursing home in Perth. Although Mrs. Lumsden had known for days that the worker would be visiting, it appeared that no effort had been made to tidy up. In fact the entire place was littered with clothing, bed clothes, old newspapers, bottles and unwashed dishes. A number of old pots and dishes were strewn about and Mrs. Lumsden stated that the roof always leaked after two or three days of heavy rain. She volunteered that the family "got by" on Social Security payments and whatever occasional work Mr. Lumsden could find. Mr. Lumsden has not had a full time job for a very long time but does find part-time manual labour--either on the docks or in landscaping. Mrs. Lumsden said that they had once tried to get a council house but had never 'heard back' about it. Mr. Lumsden and one of his friends had tried to repair the roof but their efforts really had not done much to improve the situation.

Mrs. Lumsden was quite sure that the children's head lice had disappeared and stated that Rose, Danny and Susan were all doing quite well at school. When this was queried, she allowed that they did sometimes have to miss school because of illness but suggested that all children have to do so from time to time. After further questioning, she admitted some concern over Rose; that Rose had been ill a lot and that she sometimes did not come home for a few days. Mrs. Lumsden had been trying to get her husband to talk with Rose but he kept saying that it was not up to him. Whenever

Mrs. Lumsden said anything, Rose would get angry or sulky and leave the house. She allowed that she could not talk with Rose very well but admitted no problems with any of the other children. Apparently Pat now lives in London and John is "away". Mrs. Lumsden stated that she did not know where he was. She thought that the two younger children were getting on quite well. With regard to Rose, Mrs. Lumsden suggested that it might be a good idea if someone would have a talk with her because sometimes Rose was "very strong willed" and "hard to talk to". She hinted at a fear that Rose might become pregnant but, when questioned directly, could not or would not say anything further. She said Rose was not on the pill.

Assessment:

The Lumsden's house is in a very bad state of repair. In fact one wonders that they do little about it. Even just a general cleaning would help considerably. Mrs. Lumsden, however, does not impress one as a very effectual sort of person. The impression is that she simply has very little concern about anything. Furthermore, it is felt that she is of rather limited intelligence. The slight concern which she does show over Rose has not been acted upon. Even when she suggested that the worker might speak privately with Rose, there was very little conviction to her request and the worker could not take it up because Rose had 'gone to the chemist's.' Any suggestion of other problems is met with simple denial. Mrs. Lumsden is not, however, at all hostile. The worker found herself unable to decide whether she was really ignorant of her children's apparent problems or if she just consistently used denial as a self-protective device.

It would appear that both Rose and Danny are headed for trouble. Both of them are quite young for the sorts of behavior they have been exhibiting. Danny hangs out with the local toughs, probably sniffs glue, and is often absent from school. Rose is probably sexually promiscuous, is said to drink a great deal, and quite often disappears for a few days. As far as one can tell, their parents are not particularly concerned about the children's welfare--or are, at least, inadequate parents. It is difficult to know where "ignorance and a carefree attitude" leave off and child neglect begins. In this family, however, there certainly are indications that more serious problems will arise if the situation is left to deteriorate. The worker is, however, in a quandry over where intervention might begin.

CASE STUDY

SURNAME: WELLS

MAN: Bernard, aged 57, used car salesman.

WOMAN: Tilda, aged 53, part-time domestic.

CHILDREN: Jennifer, aged 19, not residing with parents.
Frank, aged 15.

Initial Referral:

This family has been known to the Department for a number of years. We were heavily involved two and one half years ago when Mrs. Wells sought our help in controlling Jennifer. Jennifer is now married and living away from her parents. At the time of our earlier involvement Mr. Wells was confined to hospital for over three months with multiple injuries sustained in a road accident. His wife had expressed concern over his drinking but he denied having any problem with alcohol--except that his driving licence had been suspended for three years. We do not appear to have had any contact with the family for at least two years.

It was Mrs. Wells again who approached us for help, saying that Frank was becoming increasingly unmanageable. She stated that he had recently been to Juvenile Court on three charges of car theft but that, since it was his first offence, and since both Frank and Mr. Wells had offered assurances to the court he had only been admonished. Now, only three weeks later, the police had been around again making enquiries about Frank's whereabouts on certain days. Frank had told them that he was in school but, after they had left, admitted that he had not attended school on those days. He was frightened that the police would check with the school but had not been able to think of another alibi at the time. Mrs. Wells had become quite distressed and decided that a social worker might be able to help.

Home Visit:

Two days after the telephone referral the worker visited Mr. and Mrs. Wells and Frank at home. Frank had been charged with car theft again. The home is a relatively pleasant, well kept, council flat which seems comfortable enough. Mrs. Wells seemed a meek, quiet sort of woman who simply said that she wanted to keep Frank out of trouble. Mr. Wells seemed a very outgoing and hospitable man but the worker developed the distinct impression that he was slightly intoxicated. Frank said almost nothing. When asked a direct question he usually gave only a cursory reply or waited for his father to provide an answer. Mr. Wells stated that Frank had learned to drive over the summer holiday when he stayed

on an uncle's farm. The family, however, does not own a car because Mr. Wells is still banned from driving and, in any event 'it would be too expensive to run.' Mr. Wells has managed to retain his job as a car salesman and 'does a little work on the side, fixing up old bangers.' He was, of course, displeased about Frank's impending court appearance but said he would 'stand up for his son' as he had done previously. He felt that it was all a bit unfair on Frank because he was the only boy around who could drive and the other lads were always egging him into doing something he shouldn't. He also noted that Frank had never so much as scratched any car that he drove. When it was pointed out that Frank was not yet old enough to hold a driver's licence and that it appeared as if most of his cars had been stolen, Mr. Wells stated that he had spoken seriously with Frank and was convinced that he would never do it again.

Mrs. Wells very hesitatntly brought up her concern that Frank was unmanageable. She said that he was always out until late at night, that the school had been onto her about his irregular attendance and poor performance, and that he would not 'listen' to her at all. Mr. Wells allowed that Frank really should be working harder at school but also felt that he was a young man and need not be 'tied to his mother's apron strings.' He also mentioned that he had had a set-to with a school counsellor on this very subject. Frank just sat and took it all in.

Further Contact:

At the end of the home visit, worker arranged to see the family again in about a week's time--after the police reports were available. Mr. Wells accepted the idea but was obviously not pleased. However, on the next morning, Mrs. Wells called into the office, saying that she wished to talk privately. Without her husband she was more straightforward, stating that she could not do thing with Frank and that her husband's attitude was making things worse. She said that she knew of a number of other incidents when Frank had done something illegal but had never been found out. She had tried to talk with Mr. Wells about it but he would not hear a bad word about Frank and always ended up by saying 'boys will be boys.' She said that she was coming to her wits end because she received no support from her husband and was sure that Frank would eventually end up in gaol. She was also quite convinced that her husband would not hear of family counselling. He had been quite upset that she had asked the worker to call around and there had been a big row after the home visit. Her husband, she said was a heavy drinker and quite difficult at times. She was told that the worker would be involved whether Mr. Wells liked it or not because the Department would have to prepare a court report on Frank. It was agreed that we would continue with the appointment to see the family as a whole on the appointed date.

When the worker returned for the second home visit Frank was out. Mrs. Wells said that she had told him to stay at home but Mr. Wells said it really didn't matter and that we would be able to settle things better without him. Again, Mr. Wells appeared to have had a few drinks but was not really 'intoxicated'. This interview went much as the first, with Mrs. Wells timidly stating that there were 'problems' and Mr. Wells staunchly denying it. He felt that Frank was "high-spirited' but that he had now learned his lesson. When confronted with the disparity between his wife's feelings and his own, he said that Mrs. Wells tended to exaggerate and was always nagging at Frank. He said he couldn't really blame Frank for spending his evening's somewhere else. He would not admit to any problems, either with Frank or between Mrs. Wells and himself. Mrs. Wells argued with him for a short while but gave up when she didn't get anywhere. The worker arranged to have Frank visit the office on his own in order to get his story for the court report.

Frank arrived for his appointment about twenty minutes late. He was much more personable than at our previous meeting. He could, however, offer no explanation of his behaviour--only that he liked driving cars a lot. His illicit joy-rides had 'just happened.' He was a little fearful of what the court might do this time and swore that he was 'through with doing silly things.' The court, he hoped, would 'give him another chance' and possibly just put him on probation to the worker.

Frank mentioned that he spends most of his free time with a few mates--none of whom ever gets into any 'real trouble.' They hang around the youth club a lot because there isn't much else to do. He also takes weekly judo lessons at the community hall and sometimes helps his father to fix up cars. When we spoke about his mother he said that he wished she would 'get off his back' and his attitude toward women in general seemed rather disdainful. He thought he was pretty good at mathematics but really doesn't care for school very much and thinks it's just a waste of time 'because he wants to a motor mechanic.'

Assessment:

Frank impresses as being much more intelligent than his poor school performance would indicate. He certainly has learned (from his father?) how to disarm and pacify any opposition by smiling and adopting a jocular, friendly air. Nothing is taken very seriously. His only apparent frustration is his mother and his attitude indicates the he really doesn't pay much attention to her.

As with Mr. Wells, it is very difficult to engage Frank in any serious discussion. He is superficially friendly,

but one cannot break through to any deeper sharing of thoughts and feelings. He gives assurances that he is going to stay out of trouble but one gets the impression that he will say anything that he thinks you want him to say. It is difficult to know what might be done with him, or even what he really feels.

PARTICIPANT INFORMATION SHEET

Please be assured that information gathered from this form will be held in strict confidence and that the course evaluator will be the only person to see your particular responses.

PERSONAL

Name:

Mailing address:

Age: Sex:

EMPLOYMENT

Job title:

Employer:

Major duties and responsibilities:

1. For how long have you been doing this particular job?

2. For how long have you been employed in the field of social work?

3. What percentage of your working time is spent in:

	Percentage of time
- social work with individuals	_____
- social work with small groups	_____
- social work with large groups	_____
- supervision of staff/students	_____
- training of staff/students	_____
- administrative tasks	_____
- other	_____

UCATION

1. What formal qualifications (degrees, diplomas, certificates, etc.) do you hold?

Qualification	Year received

2. Have you ever previously attended a seminar, workshop, or course on integrated social work methods?

ECTATIONS and PREPARATION

4. Why are you attending this course?

5. Do you expect that it will benefit you in your job? If so, how?

6. Listed below are the titles of four books. Beside each one please mark the amount, if any, which you have read.

		Amount read
incus & Minaham	Social Work Practice: Model and Method	None $\frac{1}{4}$ $\frac{1}{2}$ $\frac{3}{4}$ All
oldstein	Social Work Practice: A Unitary Approach	None $\frac{1}{4}$ $\frac{1}{2}$ $\frac{3}{4}$ All
hittaker	Social Treatment: An Approach to Interpersonal Helping	None $\frac{1}{4}$ $\frac{1}{2}$ $\frac{3}{4}$ All
iddleman & Goldberg	Social Service Delivery: A Structural Approach	None $\frac{1}{4}$ $\frac{1}{2}$ $\frac{3}{4}$ All

Please be assured that your particular responses to the questions below ill be held in strict confidence. They will be seen only by the workshop valuator and will be reported in such a way as to preserve your personal nonymity. It is hoped that you will respond as completely, honestly, and houghtfully as time will allow.

ame: _____

ection I

Please indicate your response to each of the following items by placing n "X" on the appropriate scale. The scales do not all run in the same direction o please take care.

. Did you enjoy attending this workshop?

Yes, very much — No, not at all

. How much did you learn from the workshop?

Taught me very little — Taugt me a great deal

. How relevant was the workshop content to your on-the-job needs?

Very relevant — Not relevant

. How well balanced was the workshop?

Overconcentration on some things, not enough on others — Very well balanced

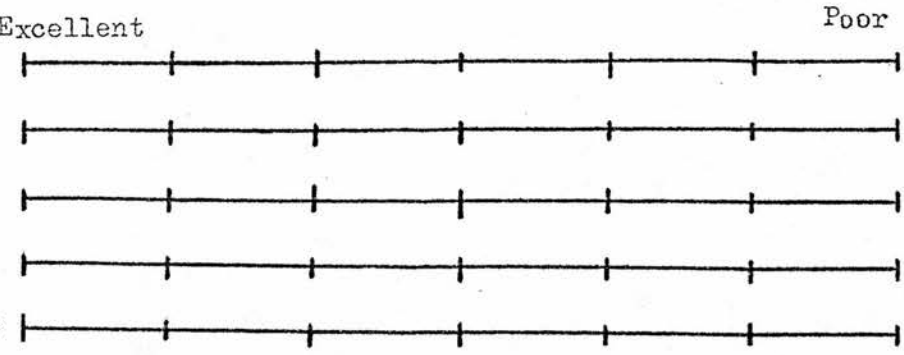
. If this workshop was offered again, would you recommend that your colleagues attend?

Would strongly recommend atterdance — Would not recommend

How adequate were the tutor's provisions for student feedback and evaluation?

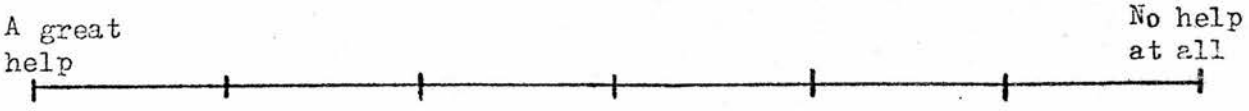
Inadequate — Excellent

. How would you rate the workshop's facilities as regards:



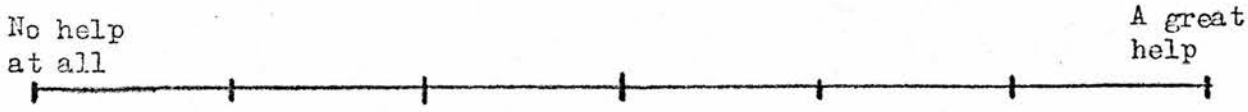
. To what extent do you think the workshop has:

(a) helped you to understand the "unitary approach"?



(b) improved your ability to apply the "unitary approach"?

(c) provided you with new ideas for agency and service reorganization?



(d) provided you with new approaches to assessment and intervention?

Section II

Please answer the following questions as completely as possible. You may use the backs of these pages, or additional sheets, if you wish.

1. What aspects of the workshop have had the most positive influence on you? How? Why?

2. Which aspects of the workshop have been most applicable to your work? How? Why?

3. Was the student "mix" appropriate;eg. the teacher-practitioner mix, the variety of work experience, the variety of agency representation,etc.? Could the mix have been improved? If so, how?

4. If you were about to organize a learning experience on the "unitary approach", how might you improve upon this workshop?

5. Have you any comments on the evaluation procedure and/or the evaluator?

6. Please use the back of this page for any further comments.

DAILY EVALUATION

Name: ..... Day: .....

I. Please rate today's sessions by placing a mark in the appropriate box after each factor.

		"Excellent" 6	5	4	3	2	"Poor" 1
(a)	Holding your interest						
(b)	Amount you learned						
(c)	Relevance to your needs						
(d)	Amount and quality of work completed in your group						
(e)	Organization, sequencing and timing of sessions						
(f)	Adequacy of tutors						

II. What was the usual participation pattern in your working group?

( ) "extremely unbalanced" – a few did most of the talking.

( ) "unbalanced" – everybody talked but some a great deal too much and others a great deal too little.

( ) "slightly unbalanced" – some talked a bit too much and others a bit too little.

( ) "balanced" – although we did not all talk the same amount, each of us talked nearly the right amount.

III. How well did your group stay on the assigned discussion topic?

( ) seldom on topic.

( ) on topic less than one half of the time.

( ) on topic more than one half of the time.

( ) on topic almost all of the time.

IV. How well did your group's tutor control the discussions?

( ) should have controlled a great deal more.

( ) should have controlled a little more.

( ) controlled just about right.

( ) should have controlled a little less.

( ) should have controlled a lot less.

V./

V. Please indicate the most important things that you learned today.

VI. Please provide any comments or suggestions on how the day might have been improved.

APPENDIX II

INSTRUMENTS FROM STUDY II

NATIONAL INSTITUTE FOR SOCIAL WORK

543

Mary Ward House, 5-7 Tavistock Place, London, WC1H 9SS (Registered Office) Tel: 01-387 9681

(Limited by guarantee: Registered No. 726022 England)

President: Lord Seebohm

Chairman: Peter M. Barclay

Principal: David Jones OBE

Honorary Treasurer: Peter E. Leslie

Registrar: David J. Pratt

4th December, 1975.

Dear

A Unitary Approach to Social Work Practice

I am writing in connection with your application to attend the above short course to be held from Monday 5th to Friday 9th January 1976, at Kingswood Hall, Royal Holloway College, Egham, Surrey.

I am pleased to be able to let you know that you have been selected for a place on the course, and I enclose an invoice for the residential fee of £28.50 which is now due. If for any reason you now find you are unable to take up the place offered to you, I should be grateful if you could let me know as soon as possible so that I may offer the vacancy to one of the applicants on our waiting list.

The staff of this workshop want you to know that they have agreed to have it independently evaluated by Ken Gordon, a post-graduate student at Edinburgh University. He is a Canadian social worker and an ex-training officer now doing research on the evaluation of social work training courses. He will be with us throughout the workshop, and will be using an evaluative method that relies on observation of the workshop and the recording of members' reactions to its various elements. He hopes that you will feel free to talk to him as the workshop proceeds; and that towards the end of the course you might find time to participate in a more formal, short, tape-recorded interview. If you can arrange it, he would also like to interview some members after the end of the course, early on the Friday afternoon. When reporting upon his observations, Ken Gordon will do his utmost to preserve the anonymity of members' responses. At the beginning of the workshop he will ask you to complete an information sheet about yourself. This and any other information that you give him will be held in strictest confidence. The National Institute is eager to collaborate in the evaluation of social work training, and we trust that you will find it interesting to participate in Ken Gordon's research.

A course programme and other details will be sent to you nearer the time, if in the meantime there is any further information you would find it helpful to have please let me know.

Yours sincerely,

Ilona A. Murdoch (Mrs.)
Course Organiser

INTERVIEW PRO FORMA

(Drafted half-way through the workshop and loosely followed in each interview)

1. What, if anything, have you taken from the workshop thus far?

2. Has anything impressed you as missing from the workshop thus far?

3. To what extent do you see the workshop as applicable to your job? Could it be more applicable? If so, how?

4. Do you think the sequencing and timing of the workshop tasks have been appropriate?

5. Would an early look at social systems theory have been a good idea?

6. Have the NISW staff been adequate to the task of assisting you in learning about unitary approaches?

7. What do you think of Kingswood Hall as a workshop environment?

8. Do you feel that the workshop has created any new needs or interests in you?

9. Has there been a particular session that you feel was especially useful or enlightening? Especially useless or dull?

10. Has the mix between theoretical and applied work been appropriate?

11. Has the mix of students been appropriate (educational and experiential backgrounds, intellectual abilities, etc.)?

12. What, if any, changes would you make in the workshop if you were the senior tutor?

13. Any comments (Impressions that have occurred during the workshop that you thought might contribute to the evaluation)?

Newfargie Cottage,
Gateside, Fife

May 9, 1976

Dear

Re: Unitary Approach Workshop - Jan.5-9/76

I am sorry that this has to be a form letter but writing each of you individually would be entirely too time-consuming. As I mentioned at the end of our workshop last January, I intend to include in my evaluation some of your reactions after having been back on the job for about four months. Now the time has come to gauge your thoughts and feelings about the workshop. Was it a useful experience? If so, in what ways has it been of value? How applicable was it to your work? Looking back on it, can you suggest ways in which it might be improved?

Would you, therefore, please complete the enclosed questionnaire (on another sheet of paper if you like) and return it to me at you earliest convenience. I am asking that you do so as soon as possible because these things tend to be forgotten if left hanging about for very long, and your "follow-up" responses are essential to completion of the evaluation. For your convenience, I have enclosed a stamped addressed envelope.

Thanks again for your co-operation and assistance in the project. I look forward to receiving your responses and hope that we might meet again one day.

Yours sincerely,

Ken Gordon
Ph.D. Candidate,
Dep't of Soc. Admin.
Edinburgh University

FOLLOW-UP QUESTIONNAIRE

Name: Position:

Please endeavour to answer the following questions as completely as you possibly can. Your responses need not be confined to the spaces provided.

1. Looking back on it, was the workshop a worthwhile experience? If so, in what ways?

2. Over the past four months, have you been able to put into practice any of the knowledge or skills you acquired at the workshop? If so, please give an account of what has happened and how your workshop experience assisted you.

3. To what extent do you think the unitary approach is realistically applicable to your work? What are the major constraints on its use in your particular situation?

4. Can you suggest ways by which the workshop might be altered in order to provide a more useful experience for future participants?